227 000

AUTOMATION OF ASTIA - 1959

A Preliminary Report

ASTIA

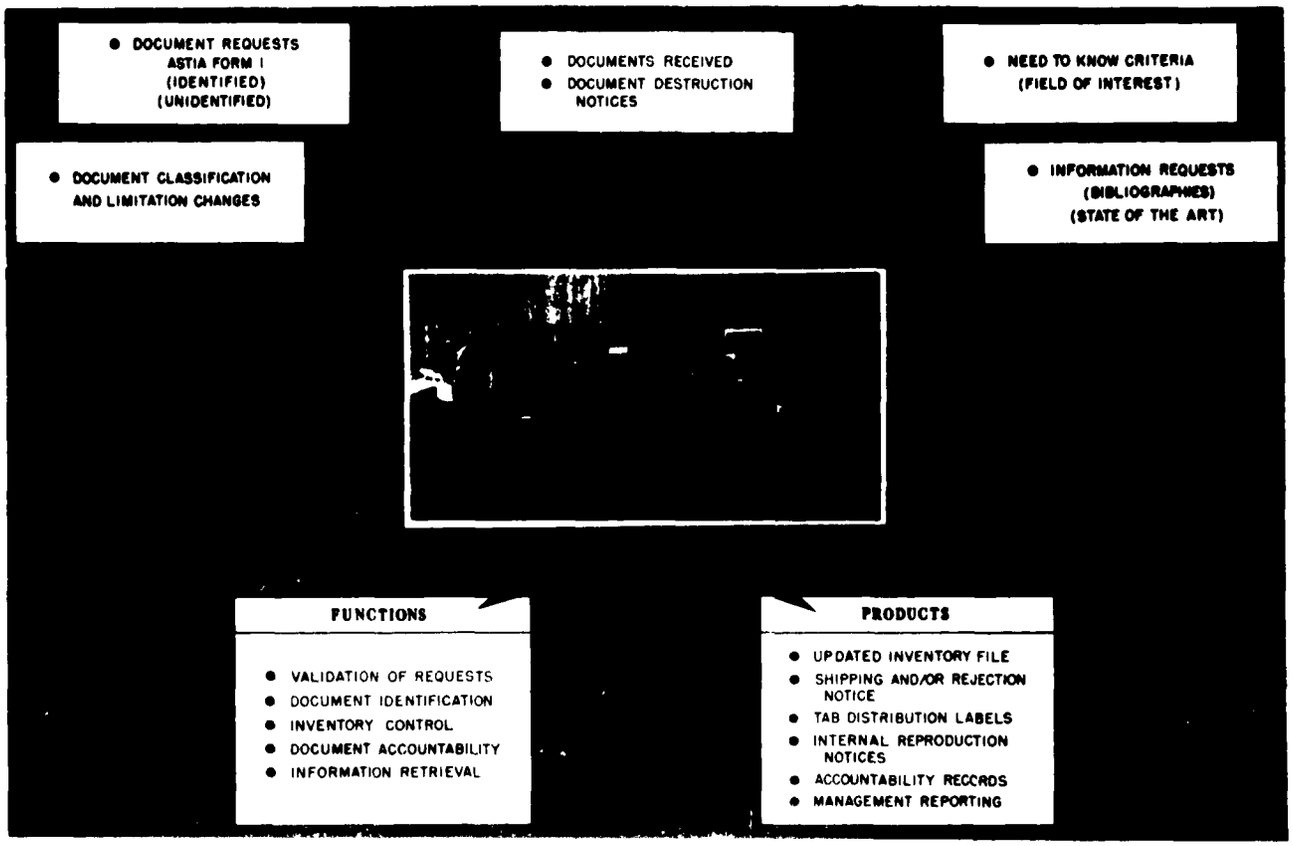
DECEMBER 1959

REPRINT APRIL 1963

ARMED SERVICES TECHNICAL INFORMATION AGENCY

• The Scientific Information Center of the Department of Defense

• An Activity of the Air Force Systems Command


• DOCUMENT REQUESTS
ASTIA FORM 1
(IDENTIFIED)
(UNIDENTIFIED)
• DOCUMENTS RECEIVED
• DOCUMENT DESTRUCTION NOTICES
• NEED TO KNOW CRITERIA
(FIELD OF INTEREST)
• DOCUMENT CLASSIFICATION AND LIMITATION CHANGES
• INFORMATION REQUESTS
(BIBLIOGRAPHIES)
(STATE OF THE ART)
FUNCTIONS
• VALIDATION OF REQUESTS
• DOCUMENT IDENTIFICATION
• INVENTORY CONTROL
• DOCUMENT ACCOUNTABILITY
• INFORMATION RETRIEVAL
PRODUCTS
• UPDATED INVENTORY FILE
• SHIPPING AND/OR REJECTION NOTICE
• TAB DISTRIBUTION LABELS
• INTERNAL REPRODUCTION NOTICES
• ACCOUNTABILITY RECORDS
• MANAGEMENT REPORTING

AUTOMATION OF ASTIA

A Preliminary Report
December 1959
AD-227 000

by

William A. Barden
Lt. Colonel William Hammond
J. Heston Heald

ARMED SERVICES TECHNICAL INFORMATION AGENCY
ARLINGTON HALL STATION
ARLINGTON 12, VIRGINIA

AUTOMATION OF ASTIA

A Preliminary Report

CONTENTS Page

FOREWORD

This is the first report to be released which describes the transition in ASTIA from a manual to an automated operation. The story here is largely one describing the planning and preparatory moves.

As this goes to press, ASTIA is at the crossroads; the point at which the phase-in of the new overlaps the phase-out of the old. Automatic Data Processing Equipment is being installed and programming under the new system is well along the way. Schedules are being met. The long used manual operations are beginning to step aside and ASTIA will soon be in operational readiness for automation.

The task of transition is an enormous one. It involves one of the largest collections of scientific and technical reports in the free world; reports that reflect the results of the major part of the United States Government research and development program during and since World War II. The task is complicated by the continuing receipt of new reports from current programs arriving at the annual rate of about 30,000 titles, and by receipt of almost 2,000 separate requests for reports from our holdings every working day. Add to this the fact that there was little precedence for such a venture, no proven system to follow, almost every step a new one to be studied and carefully planned, and the enormity of the task becomes apparent.

However, a spirit of cooperation and determination has prevailed among the people of ASTIA to the extent that not only are the automation processes being effected, but the current workloads are being met. In fact, large backlogs that once plagued ASTIA are actually melting away.

It is the long range objective to eventually, through automation, provide a comprehensive bibliographic information and announcement service in such a manner that the scientist or engineer can have at his fingertips, at any given time, information in the ASTIA collections pertinent to his needs. Also envisioned is the automation of essentially every step in filling a request for a report, including its automatic reproduction. This report relates the beginning of ASTIA's effort in these directions.

WOODROW W. DUNLOP
Colonel, USAF
Commander and Director

ABSTRACT

AD-227 000

Early considerations in automation: The history of ASTIA's experience in planning and implementing the automation of its functions is presented. Different ideas were examined and discarded in a search for a more efficient method of indexing and retrieving information. In 1953 a preliminary study based on a systems concept embracing all the functions and services of the Agency was conducted, but a full scale study was not possible until 1958. When the final selection of the Remington Rand USS-90 (Univac Solid State Computer) was made, the ASTIA staff devised methods for making optimum use of the equipment in both the business-type and information retrieval functions. Automation program: The pre-automation and automated processing of reports through ASTIA and validation of requests of military contractors is described. The three stages by which the automatic data processing system will be put into operation are examined, and the process of compiling mechanized cumulative indexes to the Technical Abstract Bulletin is presented. Creation of a Thesaurus of Scientific Descriptors: The main objectives of Project MARS (Machine Retrieval System) are: (1) to prepare a Thesaurus of descriptors; and (2) to assign these descriptors to all AD numbered reports in the ASTIA collection. The ASTIA subject headings and subdivisions were overhauled and the list reduced from 70,000 to about 7,000 headings, now termed descriptors. The scope of subject coverage was divided into about 290 generic categories called display schedules. Procedures were established for the assignment of retrieval terms, both standard descriptors from the Thesaurus and "open-ended terms" which will not appear in the Thesaurus but will provide additional retrieval access points in the form of project names, equipment nomenclature, trade names, etc.

SECTION I

EARLY CONSIDERATIONS IN AUTOMATION

William A. Barden
Chief, Research Requirements Officer

EARLY CONSIDERATIONS IN AUTOMATION

It is our objective in this paper to present a complete story of ASTIA's experience in planning and implementing the automation of its functions. That story prior to the establishment of ASTIA, actually begins when its predecessor organizations were using tabulating equipment, mechanized addressing systems, micro-cards, microfilm and continuous processors for enlarging microfilm in their efforts to provide the best possible technical information service.

In those early days (the late '40's) bibliographies were compiled for selected subjects. These were published and disseminated to organizations that were known to need the information. Important as this service was, there was a conviction that it could be vastly improved. The published bibliographies were out of date before they were distributed. Coverage in most cases was so broad as to require a considerable effort on the part of recipients to identify reports of interest to them. It was in grappling with this unsatisfactory set of circumstances that the "demand bibliography" was conceived. This type of bibliography (now called the Report* Bibliography) was created by selecting the subject headings which would bracket a user's requirements. Further refinement was attained by reviewing the abstracts to identify reports that were really pertinent to the request. Following this, the applicable catalog cards were withdrawn from the file and reproduced.

Needless to say, the users of the service received bibliographies that were as up-to-date as possible and tailor-made to their requirements. However, the methodology left much to be desired. It was strictly a manual effort and very costly. Because of this, the possibility of using punched-card techniques was carefully explored. The idea proved impractical because of limitations in reproduction methods that could be used without impairing the characteristics of the card stock. Other possibilities such as the Rapid Selector, which had been developed by Dr. Ralph Shaw at the Department of Agriculture, were considered and abandoned primarily because of the search time involved for so large a collection. To overcome this difficulty, it would have been necessary to compartmentalize the file, and this was considered impractical.

On 14 May 1951 the Secretary of Defense issued a Directive establishing the Armed Services Technical Information Agency. By the following spring a staff had been created and a "plan of operations" was being developed. At that early time it was realized that eventual automation of ASTIA's operations was essential if requirements for manning were to be held within reasonable bounds.

One of the initial steps taken by the Agency was to sponsor a contract to evaluate existing cataloging systems to select one most suitable to ASTIA's requirements. This evaluation led to the conclusion that no existing system was really suitable. Consequently, further work was undertaken which resulted in the Uniterm System of Coordinate Indexing.

---

*The term "technical report" or "report," as used in this paper, denotes that form of document utilized in recording the scientific and technical information generated through government sponsored research and development. The information in the ASTIA collections is generally recorded in this report form. Primarily, ASTIA's holdings are reports from the United States Department of Defense R&D programs, for which ASTIA serves as the central library. Generally, this report literature is considered unpublished and when supplied to ASTIA in its original form is not yet a part of the public domain.

This system was considered to offer great promise in terms of its susceptibility of automation although, initially, it was believed that the system could be implemented without the application of machines. However, it was quickly recognized that indexing and retrieval of information for a collection the size of ASTIA's required some form of mechanization, if not complete automation. In the meantime, Uniterms were being assigned to reports which were being processed into the collection. This was done to minimize the task of conversion when a machine system would ultimately be available.

Another concept that was given some consideration was the printing of catalog cards by means of embossed metal plates such as are used with conventional addressing equipment and using tabbed address plates to control distribution of the catalog cards at the time of initial printing. This idea was discarded because of the cost of developing the required equipment. Furthermore, by then, it was realized that the majority of catalog cards furnished to users were not being placed in card catalogs because of the catalog maintenance requirement.

The efforts up to this point, about 1953, had been pretty much on a piecemeal basis. The function that was the source of most of the problems at any given time was the one that was then studied with the objective of eliminating those problems. It was finally realized that a study based on a systems concept was required. Such a study must take into account all the functions performed by the Agency, plus services provided and the interrelationships among these functions and services. It was decided to sponsor a preliminary survey based on this concept. A contract was then let for a modest amount. The contractor was not in the "hardware" business, so it was believed that whatever the results of the preliminary survey, they would be valid and probably on the conservative side. The contractor surveyed the ASTIA operations and commercially available data processing equipment. The conclusions were that:

a. The request processing could be done on a small scale computer but at no saving in time or personnel.

b. The request processing could be done on a large scale computer with appreciable savings in time but at unwarranted cost.

c. Too little was known about the information-retrieval functions to warrant any action to apply data-processing equipment in that area.

The contractor recommended a full-scale study which would lead to establishing specifications of a data-processing system suitable to ASTIA's needs. However, lack of funds made it impossible to support such a study.

In the meantime, ASTIA's Management Improvement Program was beginning to pay off. Consistently improved production records were being achieved through improved methods and procedures. ASTIA staff personnel continued to keep informed on developments in the data-processing field. Progress in the development of character-reading devices was of interest because it was known that conversion of existing files would be an enormous task. Likewise, progress in the development of random-access memories, high-speed printers, and solid-state circuitry was followed with anxious interest.

During this period, ASTIA had the benefit of advice and assistance from men such as: Dr. Clarence F. Ross, then of the Directorate of Aeronautical Research, Wright Air Development Center; Dr. Howard Aiken, Director, Computation Laboratory, Harvard University (then consultant to WADC) and Dr. Samuel N. Alexander, Chief, Data Processing Systems Division, National Bureau of Standards. In addition, ASTIA staff personnel

visited McCall's in Dayton, Ohio; Sears Roebuck and Company in Chicago, and other similar organizations to determine whether any of their systems or techniques could be profitably applied to ASTIA's rather unusual operation.

Throughout ASTIA's history, there had been sporadic attempts at consolidating the Agency's operating elements. In fact, one of the objectives in establishing ASTIA was such consolidation. However, no suitable space had ever been found and nothing happened until the end of Fiscal Year 1957, when it became apparent that ASTIA would consolidate its operating elements in the Washington, D. C. area. Suitable space was found at Arlington Hall Station and, on the recommendations of the committee, the Secretary of Defense directed that ASTIA move to Arlington. Plans for automation came to a standstill. The ASTIA staff had to devote all its efforts to the preparation of a plan for the move. The plan was duly presented to the Commander, Air Research and Development Command, and was approved with a target date of 1 February 1958. The move was accomplished on schedule, and the Agency, for the first time in its history, was able to set up a well-knit organization unhampered by geographically split operating elements.

This move and consolidation paved the way for another "looksee" at the possibilities of automation for ASTIA. In March 1958, the International Business Machines Corporation offered to survey ASTIA with the objective of determining whether an Automation Data Processing System (ADPS) could be profitably utilized.

In preparing for the survey, IBM was requested to consider only commercially available equipment. This was done deliberately to focus effort on the immediate objective of quickly upgrading ASTIA's capabilities.

Working closely with ASTIA personnel, the IBM representatives made their survey and in September 1958 presented their report to the Commander of ASTIA. The survey brought out the fact that, of all the functions which were considered susceptible of automation, those which accounted for the greatest share of manpower resources were the so-called "business-type" functions. This fact is extremely important. Its consideration was the predominant factor in the selection of equipment.

It meant, in over-all consideration, that the selection of equipment should be based on requirements residing in the business-type functions, <u>but</u> with due allowance for the requirements residing in the information-retrieval functions. The selection was made on this basis and the ASTIA staff went to work on devising methods for making optimum use of the equipment in both the business-type and information-retrieval functions. The story of the planning and preparation follows.

SECTION II

THE ASTIA AUTOMATION PROGRAM

Lt. Colonel William Hammond, USAF
Automation Project Officer

A. PRE-AUTOMATION OPERATION

ASTIA receives several hundred thousand copies of technical reports each year from military research and development activities. These documents are screened to eliminate those not pertinent to the ASTIA program. The remaining reports must be checked against the existing file to identify those which have been previously cataloged. Duplicates are sent to the storage section for use in filling requests, or are destroyed in accordance with the stock-control policy. There are about 750,000 separate reports in the entire ASTIA collection today. More than 30,000 new ones are being cataloged into the system each year. The AD (ASTIA Document) collection now contains the 200,000 most recently cataloged titles. These documents were contributed by more than 10,000 different corporate authors since March of 1953. It is this part of the ASTIA collection which is of primary concern here since it will be the first under automation.

Each new report processed into the system is assigned a catalog number, and then processed through the remaining steps of cataloging. In the initial stage, classified reports are brought under security control. Reports that have passed through all phases of screening and are to be retained in the system are processed next through Descriptive Cataloging where the conventional library cataloging is undertaken. Information such as date and title of report, personal author, corporate author, contract under which the document was generated, and any other identifying information or distribution limitations are entered.

After the descriptive cataloging has been accomplished, the report is then processed by ASTIA Scientific Analysis personnel. Here its content is examined and the author's abstract is reviewed to determine if it meets the suitability criteria established by ASTIA.

If an author's abstract is not suitable, it is either modified or a new one prepared. If no abstract is provided one must be written.

The next task of the analyst in processing a new report into the collection is to assign retrieval terms which have replaced both Uniterms and conventional subject headings in anticipation of automation. From six to ten terms are required to provide for cumulative subject indexing and for the depth of retrieval desired.

The criteria for selection and assignment of retrieval terms and the development of a Thesaurus of ASTIA Descriptors are discussed in Section III of this report.

At this stage of processing, the cataloging information for each document is passed to the Reproduction Division for preparation of catalog cards and for compilation of a twice-monthly Technical Abstract Bulletin.

Until the Automatic Data Processing System (ADPS) becomes operational in February 1960, the manual operation will continue to utilize the conventional card catalogs -- three-inch by five-inch cards that are filed by originating agency, corporate author, author, subject, project name, contract number, etc. By the end of Fiscal Year 1959 a total of 7,000,000 individual catalog cards were employed throughout the ASTIA operations, requiring twenty people to file and maintain. The card catalogs, together with micro-cards of documents in the ASTIA collection, are the basis for the reference service provided by ASTIA.

The preparation of the Technical Abstract Bulletin (TAB) is geared to a twice-monthly operation for the domestic issue and once-monthly operation for a special issue for approved foreign releases.

The TAB contains the same descriptive and abstract information that appears on the catalog cards. TAB also contains subject, source, and numerical indexes.

During preparation of an issue of TAB, micro-filming of the reports to be announced in that issue is accomplished. Copies of microfilm are made to meet internal requirements for reproduction of full-size documents to satisfy requests after stocks have been exhausted. Copies of microfilm of unclassified reports that do not have distribution limitations are prepared for distribution to designated foreign countries after domestic announcement. The original microfilm negative is used to prepare a micro-card master negative for each report. Micro-cards are then produced, each containing up to sixty standard-size pages of the original report. The work flow is controlled so that the printed catalog cards and micro-cards are distributed to all catalogs in ASTIA just prior to announcement in the Technical Abstract Bulletin.

ASTIA's services are essentially limited to military activities, DOD contractors, and certain qualified government agencies. The need-to-know of a contractor for release of classified matter is compiled on the ASTIA Field-of-Interest Register (FOIR). This register is established on the basis of the 33 divisions and 242 sections of subject groupings of the ASTIA Distribution Guide. These groupings are general arrangements of the various major scientific fields. A military sponsor must establish the fields of interest of each contractor serviced by ASTIA.

A validation file is maintained which contains information on the conditions of release that apply to each report in the system. Requests by contractors for classified documents must be "validated" - checked against the approved fields of interest. All requests must be checked against any release limitations which may apply to a requested report.

Requests that cannot be validated are returned to the requester indicating the reason why the request cannot be honored. Valid requests for reports that predate the AD collection and those with AD numbers from 1 to 80,000, and any requests for microfilm are sent directly to reproduction since no shelf stock is maintained. Requests for micro-cards or for full-size documents of AD numbers above 80,000 are forwarded to the storage section for supply action. About sixty per cent of the requests received each day are for reports that must be reproduced because copies are not available in stock. This averages about 1,200 reproductions per day and represents the direct labor of 24 people.

Every day more than 200 requests are received that do not cite ASTIA catalog numbers. For each of these "unidentified" requests, the descriptive information furnished by the requester must be manually checked against the catalog cards in an attempt to identify the ASTIA catalog number. Frequently a request must be returned unfilled because it cannot be identified under the present manual operation. Those that are identified are re-entered into the request-processing system.

Requested reports coming both from stock and from reproduction are forwarded directly to the shipping section. At this point, the report is checked against the request and assembled with others for the same addressee. It is then wrapped for shipping. The shipping section prepares labels by addressograph, and maintains files of current address labels for each established user and controlling military office. Secret documents requested by contractors must be sent via the appropriate military office. In all cases involving classified documents, a copy of the request is pulled and sent to the Customer Records Section for file. This copy is held in suspense until the receipted copy is returned.

The consolidation of ASTIA in Arlington Hall in 1958 resulted in many improvements. Servicing times were brought to an optimum under manual operations. Requests for unclassified reports carried in stock are now filled in three days. It takes ten working days to fill

a request for a classified report that must be reproduced from microfilm. About three weeks are necessary to prepare an average report bibliography.

Servicing 3,000 contracts and 2,000 military customers from a stock of 750,000 documents may not sound very impressive. However, the scope of the operation can be appreciated by examining the physical activity involved in filling requests.

The requests received every day represent 1,200 to 3,500 pieces of mail that must be individually handled through the mailroom operations. Each of the requests must then be hand-sorted and processed against the master files in the Customer Services Division.

Except in rare instances, there is no way to determine which requests must be filled by reproduction until the total number has actually been checked against the shelf stock. Again, this is an individual hand operation. To reproduce the daily request workload, more than 1,000 short strips of microfilm must be hand selected from a collection of about 3/4 million such strips. The small strips must then be spliced into 100 ft. rolls for reproduction processing. In essence, one copy each of approximately 1,200 different documents must be tailor-made every day.

Each of the reports being supplied daily must be manually checked, wrapped, receipted, and addressed by the shipping personnel and distributed among 3,200 customers utilizing some 5,000 addresses. (The greater number of addresses than customers results from the fact that secret documents to contractors must be routed through the cognizant military office.)

Automation offered the only feasible way out.

Such were the conditions in ASTIA when, on 20 November 1958 the Commander of ASTIA submitted a formal proposal to the Office of the Secretary of Defense to automate a significant part of these operations.

Objectives Under Automation

The automation proposal submitted by the Commander of ASTIA was based on the following objectives:

1. Automatic request validation to determine the need-to-know and security clearance of the nonmilitary requester.
2. Automatic inventory control for improved inventory management.
3. Complete and accurate accountability for both classified and unclassified documents.
4. Mechanized index preparation for the Technical Abstract Bulletin.
5. Automatic duplication check for incoming documents.
6. Automatic identification of documents requested without reference to specific ASTIA catalog numbers.
7. Mechanized reference and bibliography service (by list of catalog numbers).

In addition to the above objectives in the original automation proposal, ASTIA has moved well along the way in formulating plans to employ punched-paper-tape equipment to prepare reproduction copy for TAB. In this manner the complete cataloging data and abstracts will be obtained in machinable form as a by-product.

Consideration is also being given to preparing punched-paper-tape for all of the cataloging and abstract information for the entire AD collection.

Somewhat longer-range possibilities include automation of the entire reproduction assembly line.

B. PROGRESS TOWARD AUTOMATION

On 15 December 1958, about three weeks after the proposal to automate ASTIA was submitted, an EDPS Planning Group composed of the following was established:

Member	Alternate
Lt Colonel W. Hammond (Chairman) EDPS Project Officer	
Mr. W. A. Barden (Tech Advisor) Chief, Research Requirements	
Mr. L. R. Barnes Chief, Management Division	Mr. H. C. Ullmann Management Division
Lt Colonel J. Feeley Chief, Support Division	
Mr. J. H. Heald Chief, Document Processing Division	Mr. H. Rehbock Document Processing Division
Mr. R. H. Chapman Chief, Reproduction Division	Mr. R. E. Reedy Reproduction Division
Mr. S. E. Pope Chief, Customer Service Division	Mr. B. C. Moyer Customer Service Division

This group first undertook the task of establishing a firm planning base. A complete review of the proposal was made, and actions required by ASTIA to attain a state of readiness for automation were identified. An all out effort was made to clean out, insofar as possible, the backlog of requests on hand, and reports to be processed into the collection.

Although ASTIA first recommended the IBM 650 RAMAC, serious consideration continued to be given to the selection of equipment best suited for ASTIA's operation. It was not until 18 May 1959 that the decision was made by ASTIA to employ the Remington Rand 90 Column Solid State Computer which had become available subsequent to the submission of the proposal. The planning for this computer was developed around a punched card configuration initially, with subsequent conversion to magnetic tape, later to be supplemented by Randex.

On 20 May 1959, a work statement and purchase request for conversion of the files, other than for information retrieval, were furnished the contracting officer at ARDC. By 30 June 1959, the initial coding of the ASTIA files scheduled for conversion to machinable form had been completed. The file conversion contract for approximately $92,000 was let prior to the close of business for the Fiscal Year 1959. The funds were made available by reprogramming within ASTIA's authorized funding.

The USAF approval, dated 24 June 1959, was received on 7 July. The Data Processing Branch with Lt. Colonel William Hammond as branch chief was immediately established under the Management Division by ASTIA General Order.

Personnel requirements for the Branch including manning, grade structure, training requirements, and reporting dates had been previously developed (see Figure 19). ARDC authorized an overage during Fiscal Year 1960 to provide for hiring of 16 civilian personnel - one short of the total requirement of the current plan.

Mr. John Meech, formerly with Statistical Services Hq USAF, was hired as deputy Branch Chief. Other employees, with the exception of one programmer and three key punch operators, were recruited from within ASTIA. Three key punch operators have been selected from among the Remington Rand contract personnel who are converting ASTIA files for input to the automated operation.

The punched-card forms, except for the user request form, were designed as a part of the conversion contract. The punched-card request form was designed by ASTIA utilizing, insofar as possible, the format and layout of the newly approved single order request form (ASTIA Form 1). The design of the punched-card request form allows it to be introduced directly into the machine process, not only for the card configuration, but in the subsequent tape and Randex configurations as well. Samples of the punched-card designs are shown in Figures 13 through 18.

Detailed flow diagrams for the complete operation under card, tape and Randex configuration to accomplish request processing, inventory control, and processing new reports into the collection were completed in early July.

The programs for the three primary daily card runs have been completed. Run one has been machine debugged. The remaining two runs will be debugged prior to 1 December 1959.

The Remington Rand Service Bureau was awarded the contract for converting the ASTIA files into machinable form. Work began on schedule and has progressed on schedule. The four card files have been created for the card operation.

File	No. of Cards
Field-of-Interest-Register	5,000
User Contract - User Code	3,200
ASTIA Master Inventory (less physical inventory data)	200,000
User and Via Address File	12,000

One additional file, the Master Index, is included in the contract. This requires the greatest effort and, by November 1959, was 60% completed. This file will not be utilized in the card operation but will be introduced directly into tape and utilized for duplication

check of incoming reports, and for identification of reports requested by other than ASTIA catalog number.

A physical inventory of the stock of AD numbered reports in storage is also included in the contract. The inventory of the classified holdings has been completed. The unclassified holdings will not be inventoried until just prior to operation to minimize the manual effort required to maintain the inventory.

Since the Fiscal Year ended prior to establishing contracts for site modification and the remaining file conversion for information retrieval, a justification for additional funds to be obligated during Fiscal Year 1960 was necessary. This amounted to $108,000 - $83,000 for file conversion and $25,000 for site preparation. The equipment layout and the specifications for the site modifications have been submitted to the Arlington Hall Station Post Engineer with the requirement that the site be ready for equipment installation no later than 1 December 1959. The Post Engineer accepted this date and began the modification on 21 September 1959 with prospects of completing the work on schedule. The air conditioning capacity for the primary source and the power capacity are already in place. Figure 20 shows site plan and equipment layout. Figure 21 shows the Univac Solid State 90 Computer with magnetic tape.

The most frequent advice ASTIA received from those who had gone through the transition from a manual to an automated operation was that the programming and data clean-up were usually grossly underestimated. They also advised that due to these deficiencies, there has never been a smooth transition into an automated operation.

The early establishment of firm automation objectives and the considerations in selection of equipment to accomplish these objectives have helped ASTIA to avoid the pitfalls of programming deficiencies. Data clean-up is another story.

Under the current program only the 200,000 titles in the AD-numbered collection are included in the automated operation. Date-wise, this covers the reports received in ASTIA since early 1953. They were cataloged under various policies, and in many instances, procedures that were not rigidly controlled. The processing operation for nearly five years was split between the ASTIA in-house operation in Dayton and the Library of Congress in Washington.

The fact that the scope and extent of the data clean-up task had been duly recognized several months in advance of actual automated operations, did not reduce the effort required to accomplish the job. Insofar as possible the data clean-up has been accomplished along with the conversion of the files. The task still remaining for the card operation alone, will require four people full time through 15 February 1960. A good deal of effort is required to standardize information to be handled by the computer. It was found that under the manual operation a single military office had as many as three address codes for the identical office symbol of a major military command.

Addresses were carried under such variations in format that it was necessary to assign one employee full time for six weeks to work out acceptable standardization.

In addition to standardization of file data, the usual number of errors of omission and commission have been spotted and corrected - in many instances only after prolonged investigation.

The major tasks of data clean-up for the card configuration should be completed well in advance of operations.

A certain amount of data clean-up and further refinement in standardization will necessarily continue for some time after the data processing system becomes operational.

C. DESCRIPTION OF THE AUTOMATIC DATA PROCESSING SYSTEM

The data-processing system to be installed by ASTIA was designed to accomplish the objectives outlined in the previous chapter--particularly the immediate objective of substantially improving timeliness of customer servicing. The system, as it is presently designed, will be placed into operation in three stages.

Stages:

Stage 1 - Punched-Card Input and Output Utilizing the Remington Rand USS-90 (Univac Solid State) Computer.

This stage will become operational 15 February 1960 and will provide the following capabilities:

Automatic Request Validation for reports in the AD collection

Automatic Inventory Control and Inventory Management Statistics

Automatic Production Control

Complete Report Accountability

Mechanized Index Preparation for TAB

The punched-card operation will employ three daily runs to accomplish request processing, inventory control and updating files:

Daily Card Run I:

Run I involves comparing identified user requests against the Field-of-Interest Master File to ascertain that a need-to-know has been previously established in ASTIA by the requester. The run also transfers the user eligibility information contained in the Field-of-Interest Master File to internal processing cards. This information is used in computer Run II when the requester's field of interest is compared to the applicable release criteria for the report requested. See Figure 1 for diagrammatic chart of Run I.

Daily Card Run II:

In Run II the identified requests and the internal processing cards from the previous run, together with their related ASTIA Inventory Master File Cards, are processed and under program control the following actions are taken:

Verifies that classified reports requested are within the requester's Field-of-Interest.

Determines report availability in stock and punches Internal Reproduction Notice, if required.

Punches updated inventory master cards.

Prints shipping and rejection notices.

Accumulates usage data on requested documents for management reporting.

Determines minimum stock levels.

Segregates output cards for subsequent processing.

See Figure 2 for diagrammatic chart of Run II.

Daily Card Run III:

Run III involves transferring the report release criteria and security classification contained in the Inventory Master File into the Validated Request and internal processing cards. Classified requests are segregated from the unclassified, and both groups are sorted to catalog-number sequence to facilitate the remaining manual operation.

Under all three configurations, numerical indexes will be included in the twice-monthly TAB. Cumulative subject and source indexes will be compiled quarterly. The fourth quarterly cumulation each year will be in the form of an annual cumulative index for the entire year. See Figure 3 for diagrammatic chart of Run III.

Stage 2 - Punched-Card-Magnetic Tape Input and Output Utilizing USS-90 Computer. Magnetic tapes will be installed on 1 July 1960. In addition to incorporating the functions of the card operation, the tape system will also provide the following capabilities for handling the AD collection:

Automatic identification of reports that are requested without reference to specific ASTIA catalog numbers.

Automatic duplication check for incoming reports.

Mechanized bibliography compilation.

Mechanized information retrieval based on retrieval-term search techniques.

The tape system will employ both punched-card and magnetic tape input-output.

The operations accomplished on the three daily card runs will require only two daily tape runs. An additional daily run will accomplish duplication check and identification of reports that are requested without reference to ASTIA catalog numbers. A separate program is being designed for bibliography compilation and information retrieval.

Initially, in the tape operation, identification and duplicate checking will be accomplished only for the AD collection. As soon as a usage pattern can be developed for reports in the older collection, records for these documents that are being requested frequently will be added to the master tape files. See Figures 4, 5, and 6 for diagrammatic charts of the three daily tape runs.

Stage 3 - Punched-Card-Magnetic Tape Input and Output Plus Randex Operation
In the punched-card configuration, request processing and inventory control require three separate machine runs. The elapsed time from the beginning of these runs until valid re-

quests can be filled is approximately 7 hours. Inventory control involves maintaining data on activity of specific reports for use in computing reorder point and reproduction quantity, and for listing those specified for destruction.

In the tape configuration these same functions plus identification and duplication check will be accomplished in three separate machine runs. The elapsed time is estimated at about five hours.

In the Randex configuration the functions performed in the five hours of tape operation will be accomplished in two machine runs in four hours elapsed time.

The principal advantage in the use of magnetic tapes is that the sorting, merging and searching functions can be accomplished much faster than with punched cards; while the principal advantage in the use of Randex is that the random access feature eliminates the sorting and merging functions and reduces the searching to a machine look-up of the address at which specific information is stored.

In the Randex configuration request-processing time is so reduced that considerable computer time can be devoted to subject searches and bibliography compilation. The reason for employing tape for this information-retrieval process is that the subject search normally will involve the entire AD collection, and in such a situation the search can actually be accomplished faster with conventional tape units than with the random-access device which is at its best when only parts of a file are involved and even then when the identity of the desired information is already known. Obviously in the field of information-retrieval this is not the case. Figure 7 shows the general system diagram under card operation while Figure 8 shows the system under magnetic-tape-Randex operation.

Incoming Report Processing

Under the automated system the present basic flow pattern for the cataloging and analysis of incoming reports will remain. However, the descriptive cataloging information input for each report received will be done on the same daily run that is made for identification of reports requested without AD numbers.

In the tape configuration, the descriptive information for reports received and that for identification of requested reports will be converted directly to magnetic tape on a Unityper and compared against similar information in the tape Master Index File. To reduce the total amount of tape time required and to establish uniformity of listing, a numerical code has been assigned to each corporate author.

If the duplication check reveals that the report has already been cataloged into the ASTIA collection, the computer will indicate the correct catalog number for the report and the dispositions to be made of the additional copies. The inventory level will be automatically adjusted.

If the check indicates that the report is a new title, the computer will assign the appropriate catalog number. A temporary record card of this new title will be retained until the cataloging process has been completed. The completed records will then be entered into the master ADPS files.

As explained in Section III scientific report analysts will provide retrieval terms and their code numbers for programming. Retrieval terms to be used for cumulative indexes to TAB will be identified for subsequent machine operations. Until Unitypers are available for direct application of magnetic tape, a punched-card will be prepared for each retrieval term to update the file. Terms that are to be utilized for the indexes will be punched in both alpha and numeric code. Other retrieval terms will be punched in numeric code only.

As mentioned earlier, preparation of the reproduction copy for TAB and catalog cards will be greatly simplified by employing a punched paper-tape device, which creates a paper-tape record of the information being typed, together with corrections, etc. The paper tape is read back through a typewriter to create the finished copy for TAB and catalog cards. It can also be utilized to create magnetic tapes for subsequent use on the high-speed printer to print out bibliographies or to produce a mat for further reproduction. The application of this phase is under development.

The degree to which abstracts can be accommodated within the character limitations of currently available equipment presents a challenging problem. Equipment manufacturers, particularly Friden, Remington Rand Univac, IBM, and RCA have this aspect of the problem under active consideration. As an interim measure, ASTIA will minimize the use of special characters in abstracts. In the meantime, ASTIA is earnestly seeking usable methods for recording special characters commonly used for expressing mathematical and chemical formulas.

See Figure 9 for flow chart of document processing under punched-card operation and Figure 10 under the tape-Randex operation.

Request Processing

Servicing of requests for reports in the ASTIA system must be rapid, and economical, and since it involves release of classified information, it must be accurate.

To achieve these goals, a rather novel design of the 90-column Remington Rand punched-card was devised to mechanize the FOIR data. A compatible design was incorporated into the punched-card designs to mechanize inventory management data and to replace the present request form. The punched-card request form will also continue to serve as an input medium to the tape and Randex configurations of the computer.

User Agency Codes for all military users and holders of single contracts will be pre-punched into the request card forms prior to distribution to users. Only the identifying address code will be pre-punched for multi-contract users. A minimum of information must be inserted by the user. For a request for a specific report identified by ASTIA document number, the user need only check the type of copy (full size, micro-card, microfilm) and write the ASTIA number. Contractors are also asked to cite the complete user code number plus the contract number on which the request is based. This will serve as a cross reference to minimize errors and to insure positive identification by a via or certifying office.

Once received at ASTIA, all request cards are processed through the various steps shown on the chart in Figure 11 under the punched-card operation and in Figure 12 under

the tape operation. The operation and computer programs under the tape and Randex configurations are similar in nature to the card operation, but also include identification of documents not requested by ASTIA document number.

If a report requested cannot be identified or its release is not authorized by the requester's FOIR, a high-speed-printer form letter will be addressed to the requester indicating the reason why the request cannot be filled.

Report Control and Accountability

Accountability and control of all ASTIA reports, internally and externally, is an integral part of the ADP System. The present manual system under which complete records are maintained for classified reports only, will be expanded to include the maintenance of a complete record of all transactions. This will permit machine processing of actions involving classification upgrading, limitation changes, recall because of technical errors, and recall upon contract termination.

In the final stage of Daily Card Run III, validated user requests are separated into classified and unclassified groups. Each request for a report will be supported by a FOIR card form containing the release criteria applicable to the requester and the report requested. The FOIR card form will be used internally by ASTIA as a record of the transaction and for accountability record in the case of classified reports.

As classified material is prepared for shipment, a punched-card request form which also serves as a receipt form, will be placed inside the package. The punched-card form used internally for machine processing of requests will serve as an outgoing security log.

When the user request card is returned receipted, it will be machine processed against the outgoing log maintained in punched-card form. The matched accountability card will be transferred to the permanent accountability file.

The classified outgoing log will be reviewed automatically by the computer or a collator matching process to determine delinquency.

Appropriate follow-up forms will be created where required to remind users of their delinquency in returning receipts for classified reports. Unclassified reports do not require suspense action. Therefore, upon shipment, the user request card is placed immediately in the accountability file.

Inventory Control and Reproduction

The most economical approach to reducing the request-to-response time within ASTIA is to increase the ratio of requests which can be filled directly from the shelf. Several actions have recently been taken toward this end. Military instructions have been revised to require contributors to supply ASTIA with ten copies of all newly published technical reports. ASTIA has instituted a procedure in the screening process to remind contributors of any delinquency in providing the ten copies required on initial distribution. Also, in many instances, it is possible during the screening process to identify reports that will have high request activity. The contributors of these popular reports are being requested to furnish additional copies, if available. These procedures, designed to increase the stock in the more active areas, have been instituted too recently to permit full evaluation of their

results. Under current operations 55% of the requests are for reports not carried in stock and must be reproduced from microfilm.

A program of perpetual inventory control and usage statistics for each report of the AD collection will be maintained under the automated operation. Reproduction will be accomplished to fill anticipated orders on an automatic analysis of the inventory usage statistics. These statistics will also be maintained for the older collections (received prior to the start of the AD numbering system in 1953). If the use pattern indicates that it is feasible to do so, limited stocking of these reports will be made.

Under the manual operation it has been possible to develop reliable usage statistics for classified reports. Full advantage has been taken of these statistics in designing the computer program for inventory control.

At the same time that selective reproduction for the shelf stock is being initiated by the computer, selective destruction of low-activity stock will also be taking place under computer control. This action is necessary to provide storage space for stocking the high request activity reports.

The reproduction work order will be machine created and sorted by catalog number, type of reproduction, number of copies, film size, and other factors to simplify production control throughout the reproduction processing. Work orders will be initiated either as a result of a request for reports for which no hard copies are available, or as a result of the analysis by the computer of the report usage statistics.

No economic justification has yet been developed to change from microfilm as the primary storage medium. Various information storage methods and mechanisms are still being investigated.

The continued use of the Haloid Xerox Copy Flo printer supplemented by the Eastman Kodak continuous wet process printer to reproduce from roll microfilm provide the best solution to ASTIA's immediate and midrange system requirements.

At the suggestion of ASTIA, the Haloid Xerox Corporation has undertaken to improve the production speed as well as the half-tone reproduction capabilities of the Copy Flo Printer. Also some discussions have been held with RCA, Magnavox, and the General Precision Corporation to explore the future possibilities of Video tape in automating both the reproduction and retrieval operations.

Based on the current program level, it is estimated that a 20% increase in the reproduction workload for a period of six months to one year will be required to carry out the planned prestocking program. In anticipation of this increased workload, an additional Haloid Xerox Copy Flo printer has been acquired. Two new high-speed Eastman Kodak continuous flow wet process printers have also been ordered to replace two older printers. An order has also been placed for two additional offset presses to provide an increased capability for multi-copy reproduction.

Bibliography and Reference Service

In the initial stages of developing a machine retrieval system for the ASTIA collection of reports, four major tasks were identified and projects for their accomplishment undertaken. The tasks were:

1. Development of an authoritative list of retrieval terms for the information contained in the AD collection.

2. Assignment of retrieval terms to the reports in the AD collection.

3. Converting assigned retrieval terms into machinable form.

4. Selection of suitable means of mechanizing the retrieval of information.

The first two of these tasks were assigned to the Document Processing Division of ASTIA. Section III of this report is devoted to accomplishments in these two areas.

The third task involving the conversion of the retrieval terms into machinable form is being accomplished under contract. Approximately 1,700,000 individual punched-cards will be prepared for the entire AD collection.

The fourth task, mechanized information retrieval, was considered a special problem. This is because data processing equipment is basically designed to work with numerical data and because serial searches through large masses of data can be extremely time-consuming and costly. The first step in solving the problem is to represent only AD numbers and retrieval code numbers in the system. Thus, the data processing equipment will be dealing with numbers in a simple matching routine. The second step in solving the problem is to compartmentalize the information in order to reduce search time to a minimum.

In the punched-card configuration it was originally planned to do extensive experimental work with conventional collators to gain operational experience. However, Remington Rand has made arrangements to deliver the complete tape configuration several months in advance of the scheduled operational date. The tape units will be utilized for extensive testing as the file conversion progresses.

Under the punched-card configuration, machine runs for search and print-out are estimated at about 30 minutes per bibliography. Under tape configuration the search and print-out will be accomplished in about 15 minutes. In both cases the output will be lists of pertinent catalog numbers. Catalog cards with complete bibliographic data may then be drawn from the catalogs.

In order to finally eliminate the need for catalog cards it will be necessary to put all the information they contain on tape. This would make possible automatic print-out of bibliographies, each reference containing the abstract and all the descriptive information. ASTIA is probably another year away from this program.

Publication of a Composite Field-of-Interest Register

ASTIA has long recognized the need for a more thorough and comprehensive initial distribution of scientific and technical reports directly by the originators of such reports. This would provide ASTIA users with more timely service and also tend to reduce the number of requests received by ASTIA for these reports.

Mechanization of the Field-of-Interest Register will enable ASTIA, at least on an annual basis, to machine compile a list of the DOD contractors, showing their fields of interest. This list can be provided as a guide in making out distribution lists on a subject basis.

Since the FOIR will be converted to machinable form several months in advance of the

installation of ASTIA ADP equipment, machine time will be made available by the Remington Rand Univac Service Bureau to compile the initial list.

If experience indicates that it would be profitable to expand the list to include other than DOD contractors, other users will be requested to advise ASTIA of the fields of interest they desire to have listed.

D. CUMULATIVE INDEXING AND TAB PREPARATION

Mechanized Cumulative Indexing

Proper utilization of the technical information store maintained by ASTIA demands both prompt announcement of availability of new information, and adequate, easily used reference keys.

The ASTIA announcement function is provided for by the Technical Abstract Bulletin issued on a twice-monthly cycle, announcing newly acquired reports on the average of 60 working days after receipt. Ultimately, under the automated system the aim is to reduce this to 20 to 25 working days.

The entries in TAB are listed in sequence by subject division of the ASTIA Distribution Guide. This in itself acts as an adequate means of discrimination to guide the user in the initial perusal of each individual bulletin. A numerical index to each issue is included for added convenience.

The subject and corporate author indexes will be eliminated from the individual issues of TAB and in their place a Quarterly Cumulative Index will be published for the first three quarters of each calendar year. At the end of the fourth quarter an annual cumulation will be published.

These cumulative indexes will increase the scope of service to the ASTIA user, and should reduce the volume of unidentified requests. The indexes will also allow the engineer or the scientist, personally if he so desires, to conveniently and easily seek out required technical information. This should also enable him to discriminate more accurately and precisely in the selection of required information.

The cumulative indexes will be compiled by subject matter and by corporate author and referenced to AD number, TAB issue, and the appropriate division in the ASTIA Distribution Guide. These indexes will be prepared from the card files prepared for input to the Retrieval Term File, Master Index, and Inventory Files. Punched-card decks used to update the Master Inventory File during each semi-monthly cycle will be retained and used to print out the numerical index for the corresponding issue of TAB. This group of cards will be gang punched with the TAB issue number. The group already contains the division and section of the Distribution Guide. The originating Agency code will be transferred from the punched-cards used to update the Master Index File. Selected groupings of the retrieval terms will be utilized for subject accumulations for indexes. The punched-cards used to update the retrieval term files for six TAB cycles (quarterly) will be run against the numerical index card decks to pick up TAB issue number and division listings. The card decks will then be utilized to print out repro copy for the cumulative indexes.

Mechanized Technical Abstract Bulletin (TAB) Compilation

A simpler procedure for providing proper and complete service to new contractors of the Department of Defense as well as those entering new fields of activity is desired. The problem is magnified in that back issues of ASTIA's announcement publications are not available for distribution. Also, prior to September 1957, the announcement publications did not carry abstracts. Thus it is difficult to satisfy new users with broadly defined technical information requirements.

The preparation of a form of cumulative TAB, at least for the AD collection (documents cataloged since 1 March 1953) offers the most practical solution to the problem.

Issuance of TAB and bibliographies prepared photographically from catalog cards and of subject and corporate author indexes covering the AD collection has been considered by ASTIA. Thus far, producing such publications in this manner has been determined to be too costly and beyond ASTIA's available resources.

As previously discussed, it is planned to employ punched-paper-tape devices to convert cataloging and abstracting data into machinable form while preparing reproduction copy for TAB. By converting the paper tape to magnetic tape, ASTIA can employ the 600 line/minute high-speed printer to produce repro copy for TAB and bibliographies at comparatively low cost. It is recognized that the difficulties imposed by the limitation of characters available in existing computers and printers must be overcome before full effectiveness can be achieved in this area.

Consideration is also being given to converting into machinable form all of the cataloging and abstract information for the entire AD collection. It is hoped that this can be completed by contract on schedule with the installation of the USS-90 Tape-Randex configuration. This would give ASTIA considerable flexibility in compiling announcement and reference information. Any selection or grouping of reports or references to reports in the AD collection could be compiled and indexed mechanically. As ASTIA gains experience in the rapid response information service, new concepts are bound to develop.

E. FIGURES

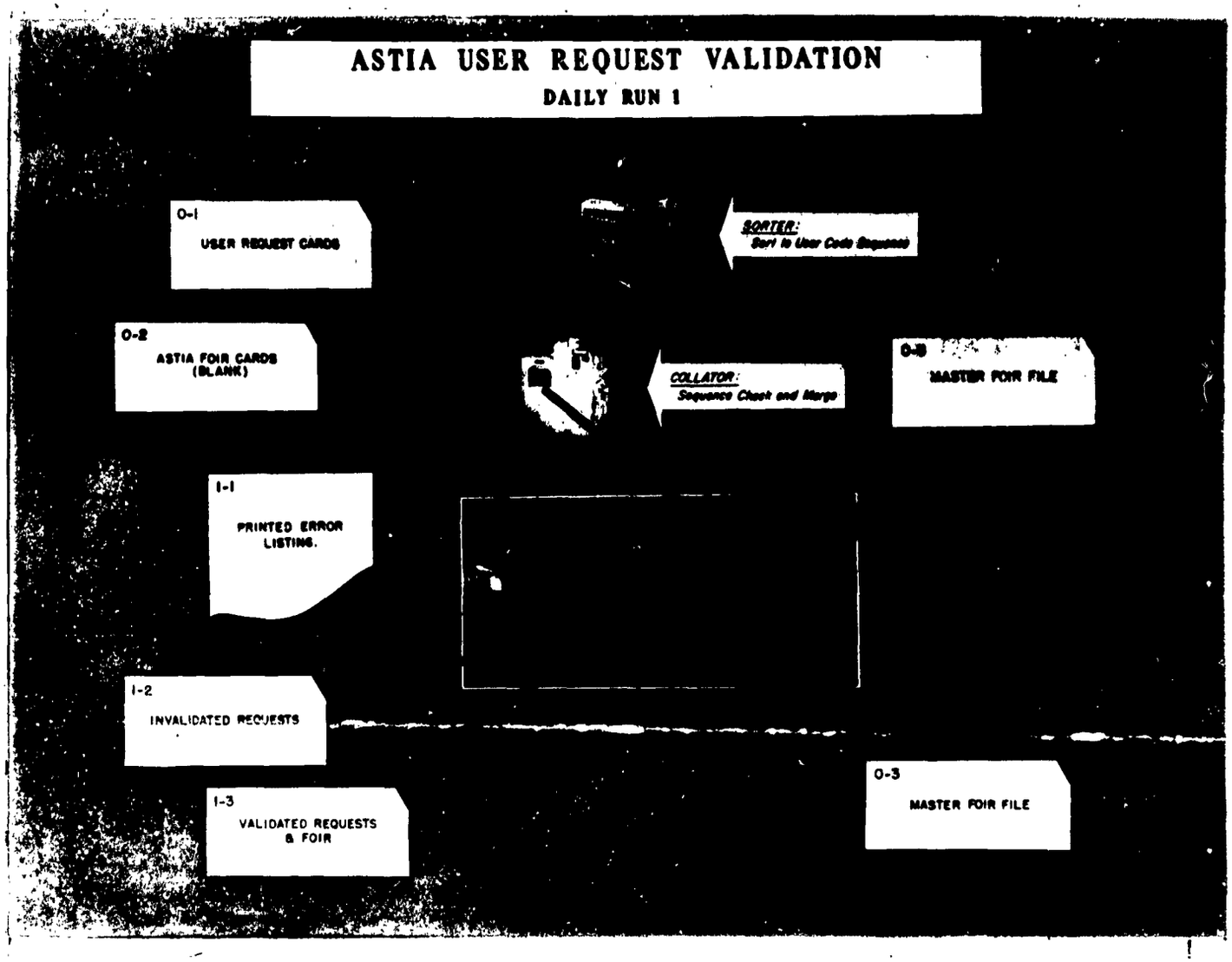


Figure 1

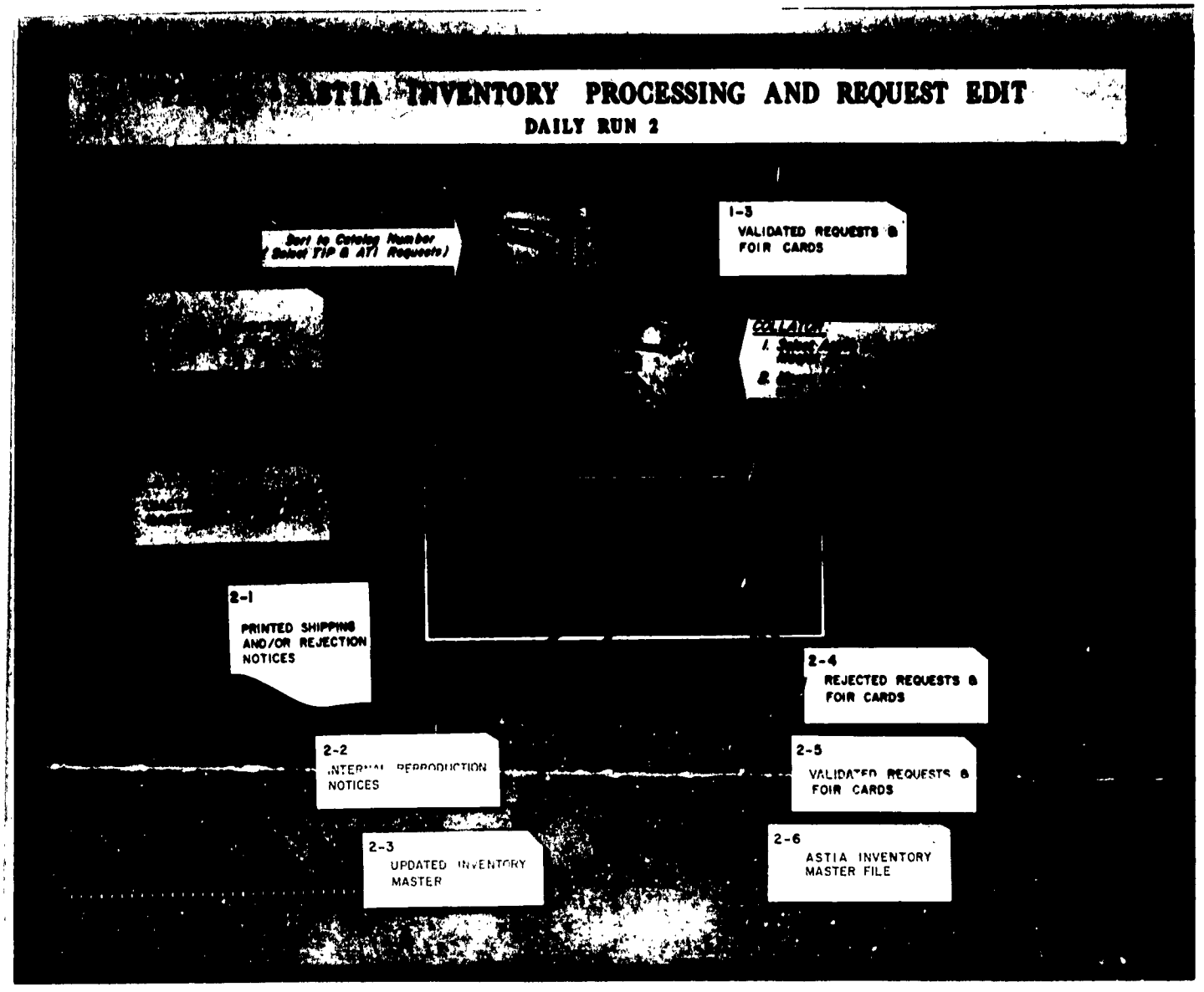

ASTIA INVENTORY PROCESSING AND REQUEST EDIT
DAILY RUN 2
Sort to Catalog Number
(Select Tip & ATI Requests)
1-3
VALIDATED REQUESTS &
FOIR CARDS
2-1
PRINTED SHIPPING
AND/OR REJECTION
NOTICES
2-4
REJECTED REQUESTS &
FOIR CARDS
2-2
NOTICES
2-5
VALIDATED REQUESTS &
FOIR CARDS
2-3
UPDATED INVENTORY
MASTER
2-6
ASTIA INVENTORY
MASTER FILE

Figure 2

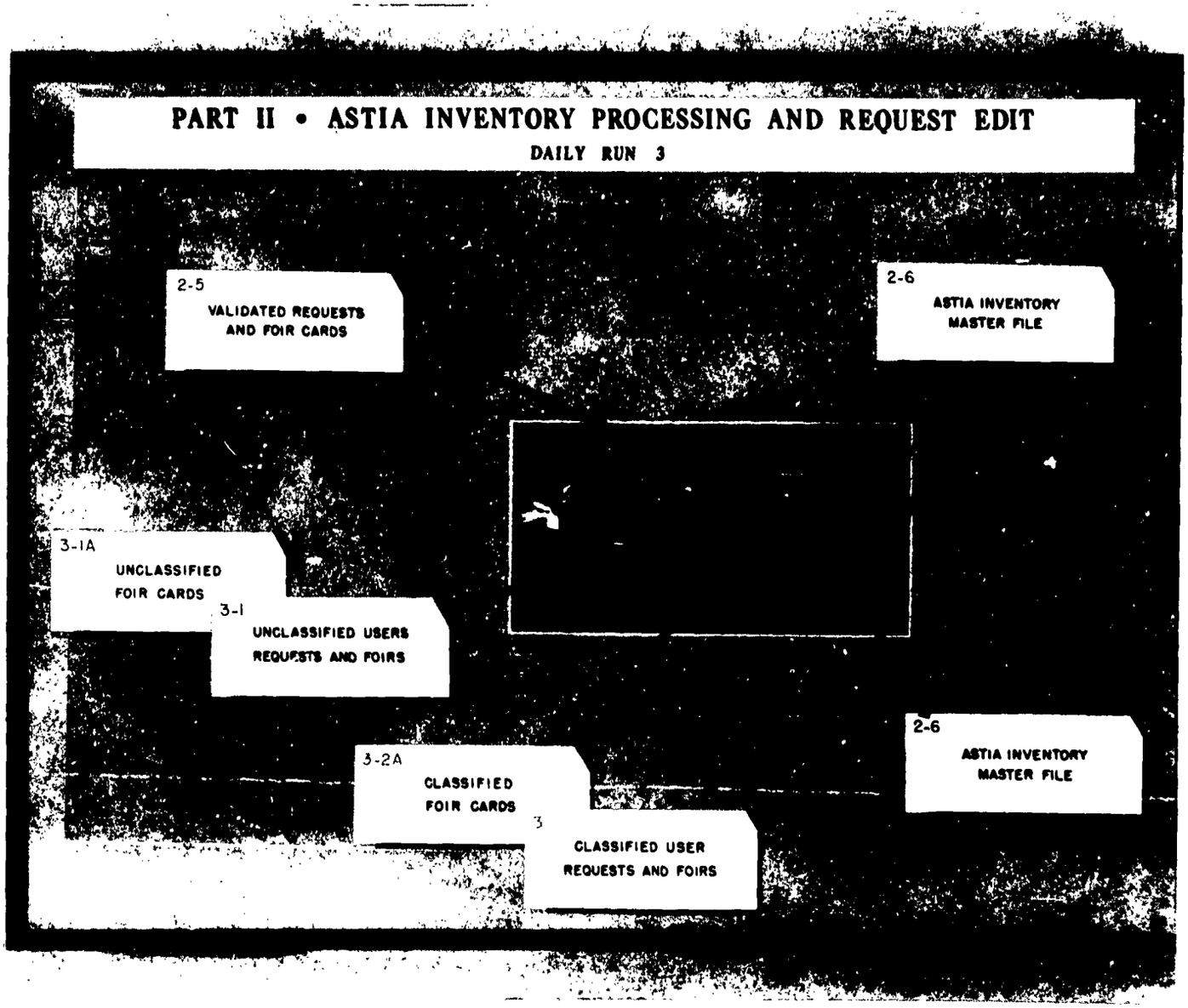

PART II • ASTIA INVENTORY PROCESSING AND REQUEST EDIT
DAILY RUN 3
2-5
VALIDATED REQUESTS AND FOIR CARDS
2-6
ASTIA INVENTORY MASTER FILE
3-1A
UNCLASSIFIED FOIR CARDS
3-1
UNCLASSIFIED USERS REQUESTS AND FOIRS
2-6
ASTIA INVENTORY MASTER FILE
3-2A
CLASSIFIED FOIR CARDS
3
CLASSIFIED USER REQUESTS AND FOIRS

Figure 3

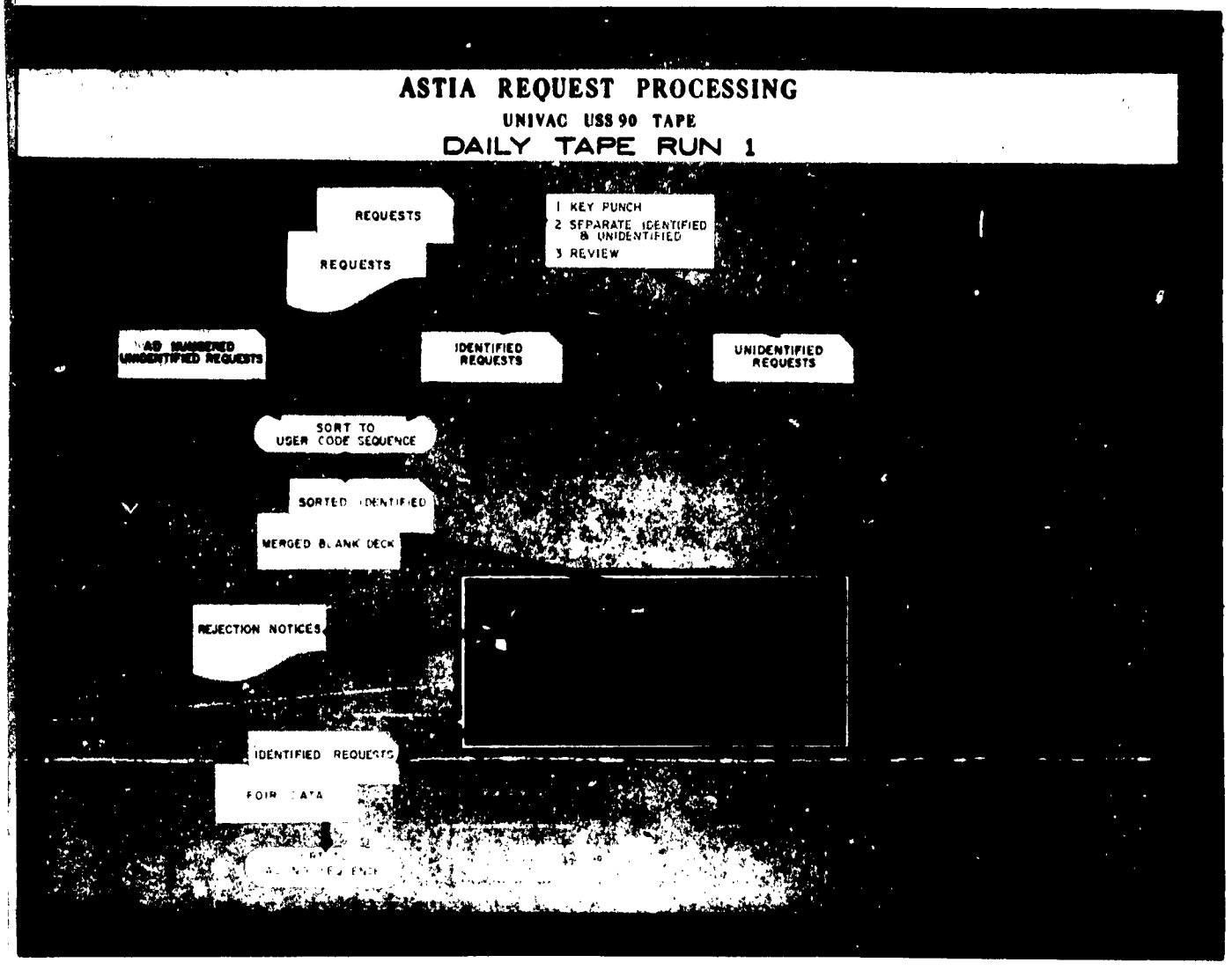

ASTIA REQUEST PROCESSING
UNIVAC USS 90 TAPE
DAILY TAPE RUN 1
REQUESTS
REQUESTS
1 KEY PUNCH
2 SEPARATE IDENTIFIED & UNIDENTIFIED
3 REVIEW
IDENTIFIED REQUESTS
UNIDENTIFIED REQUESTS
SORT TO USER CODE SEQUENCE
MERGED BLANK DECK
REJECTION NOTICES
IDENTIFIED REQUESTS

Figure 4

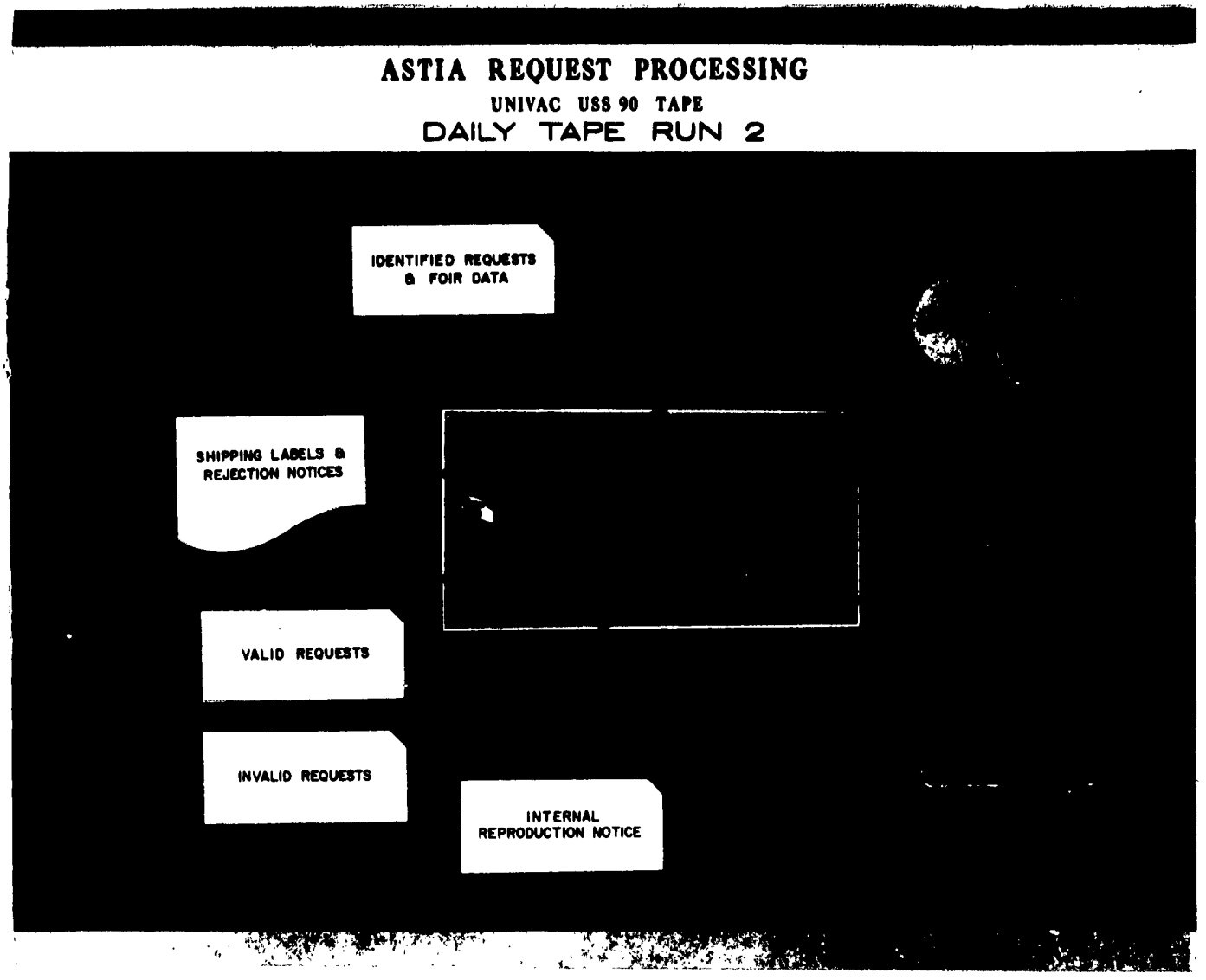

ASTIA REQUEST PROCESSING
UNIVAC USS 90 TAPE
DAILY TAPE RUN 2
IDENTIFIED REQUESTS & FOIR DATA
SHIPPING LABELS & REJECTION NOTICES
VALID REQUESTS
INVALID REQUESTS
INTERNAL REPRODUCTION NOTICE

Figure 5

ASTIA REQUEST PROCESSING

UNIVAC USS 90 TAPE

DAILY TAPE RUN 3

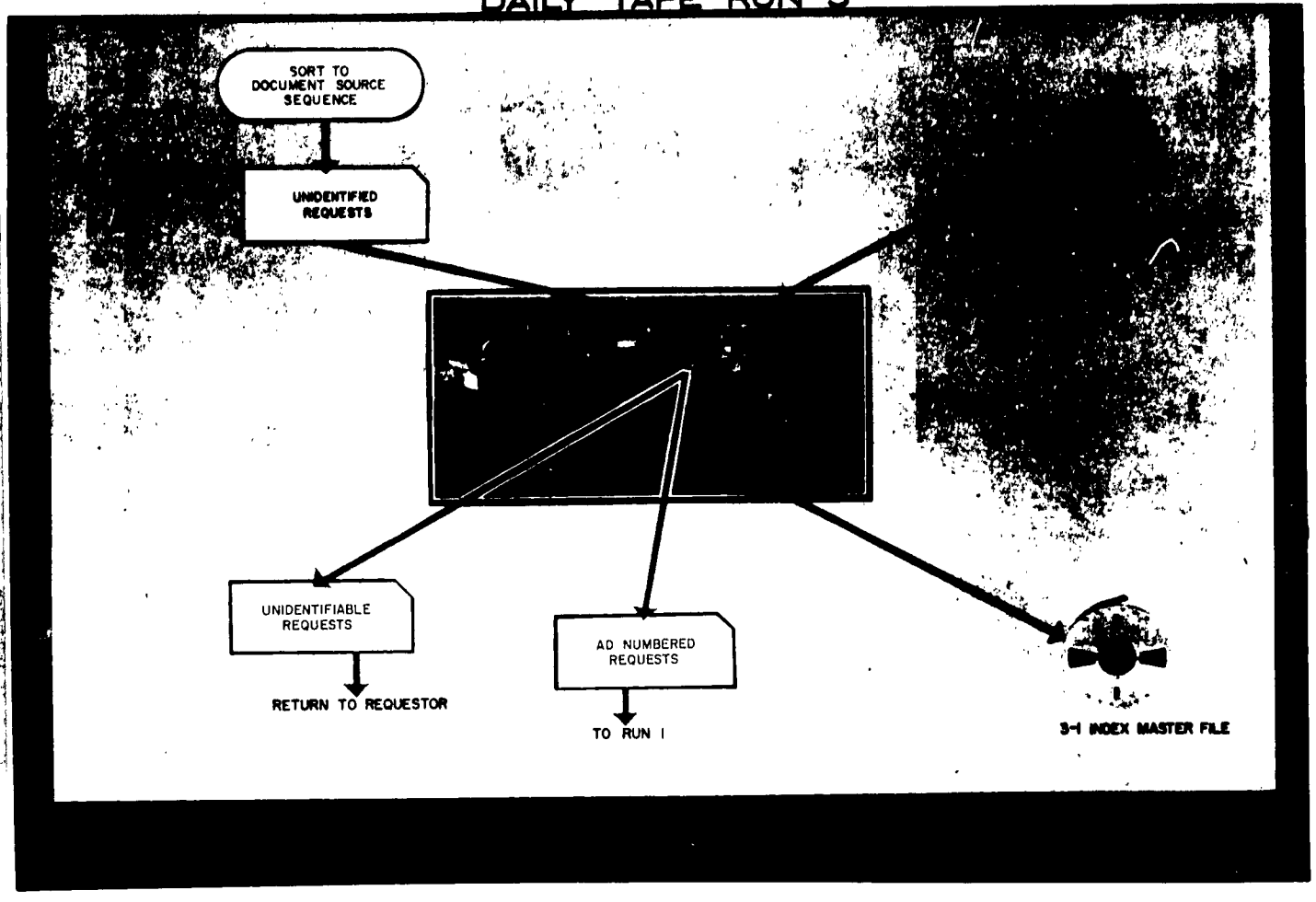


Figure 6

GENERAL SYSTEM DIAGRAM

CARD CONFIGURATION

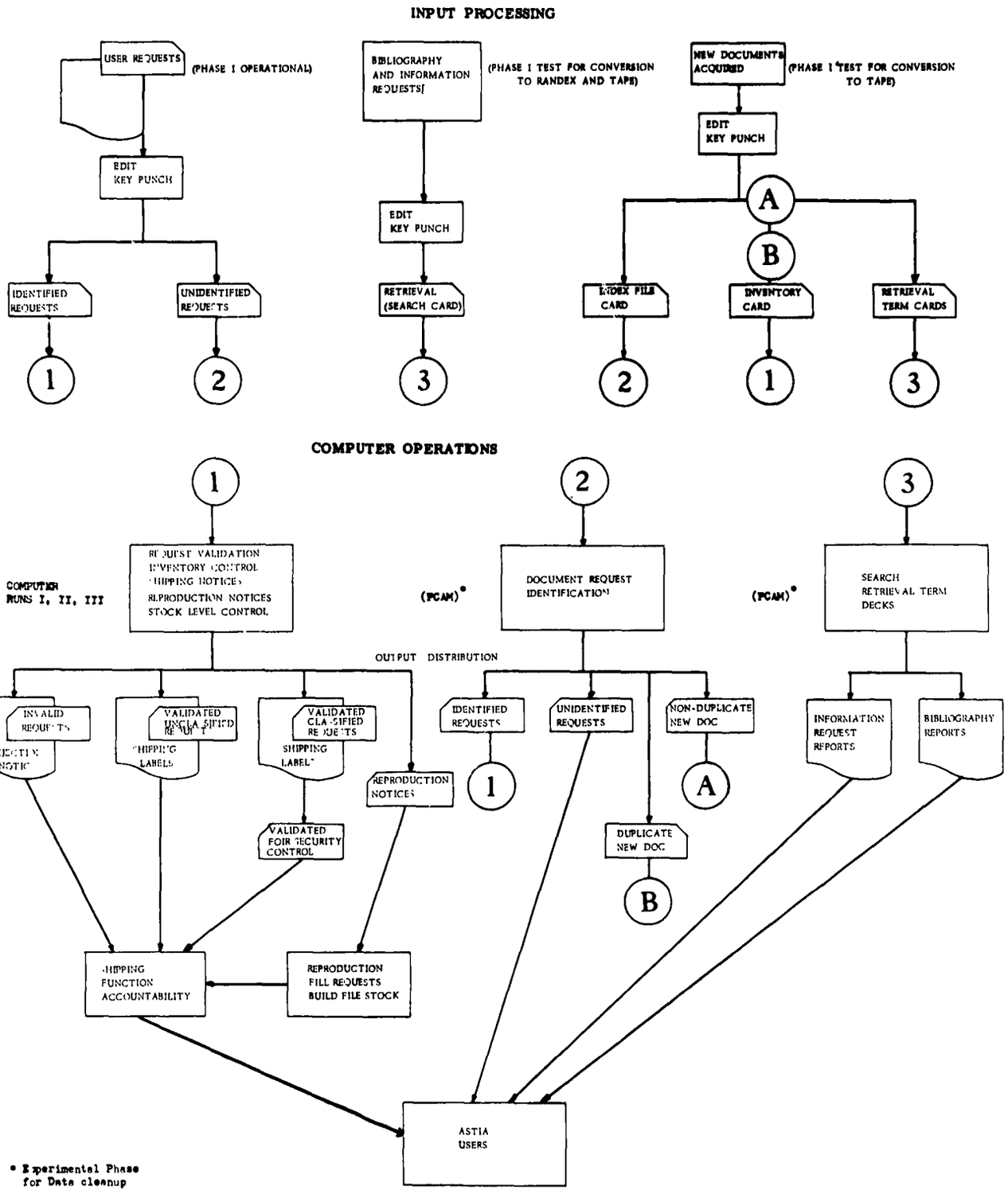


Figure 7

GENERAL SYSTEM DIAGRAM

TAPE AND RANDEX CONFIGURATION
(Phase Two and Three)

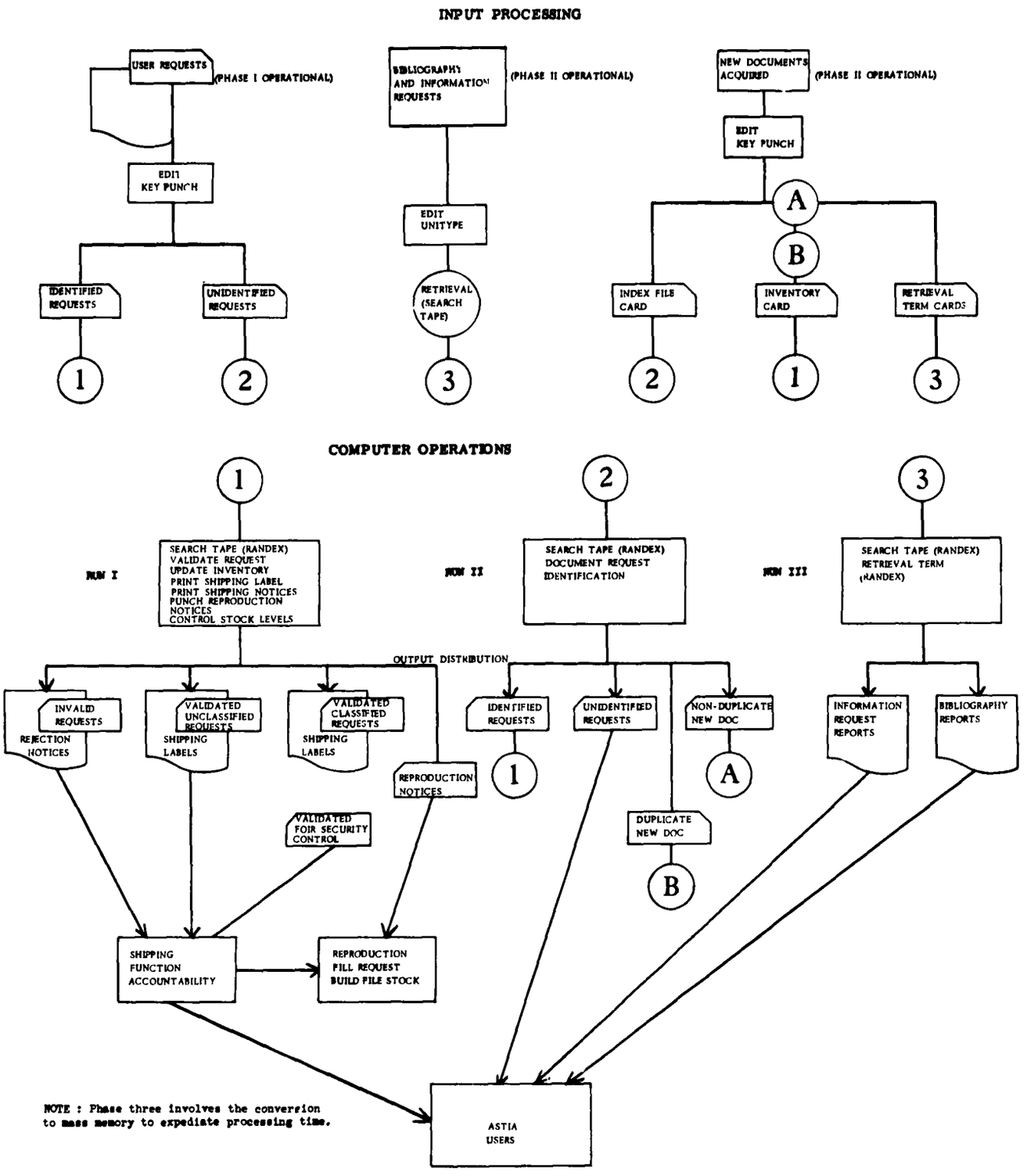


Figure 8

DOCUMENT PROCESSING

USSRO Punch Card Configuration

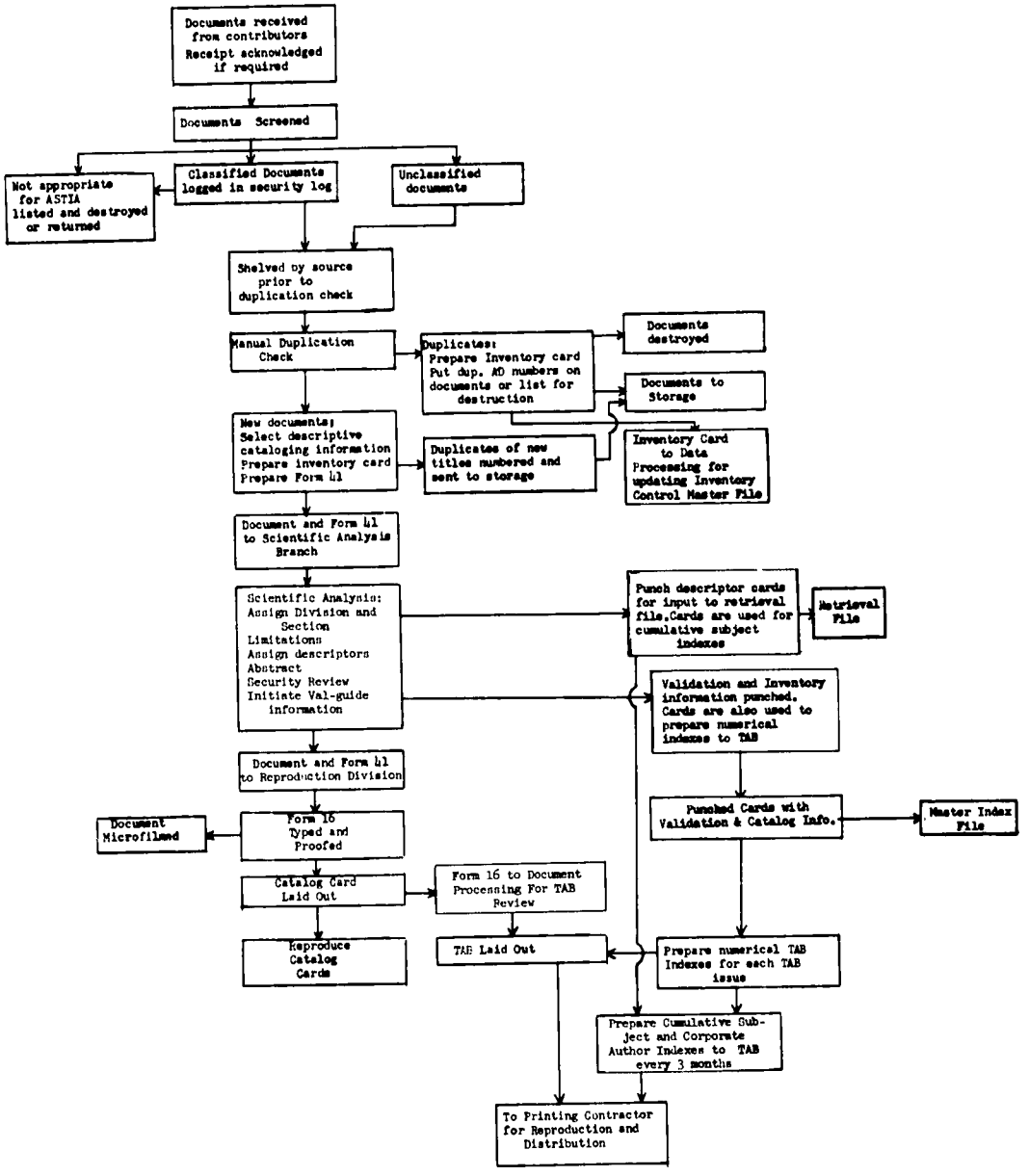


Figure 9

D O C U M E N T P R O C E S S I N G
USS 90 Tape and Randex Configuration

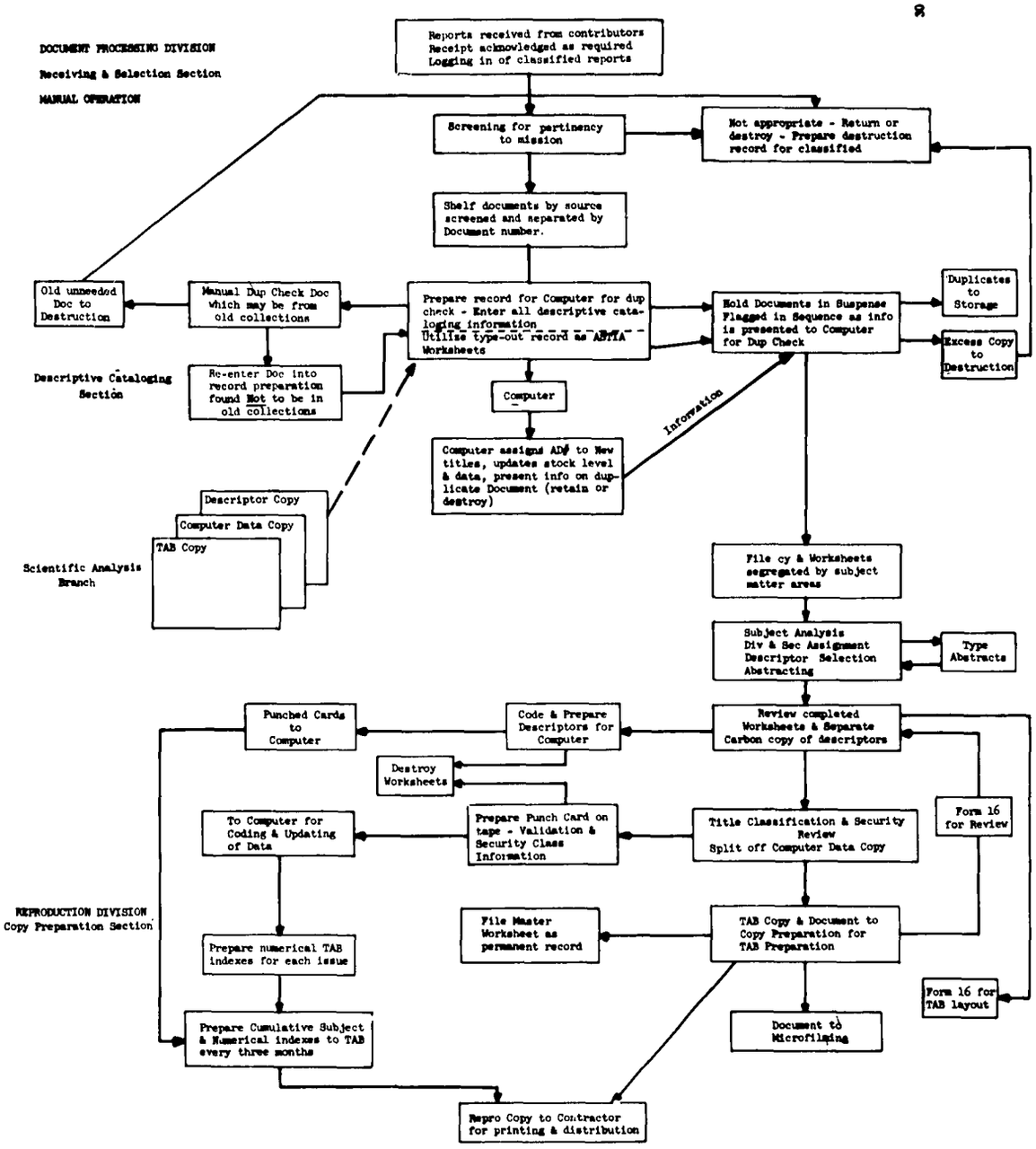


Figure 10

REQUEST PROCESSING

USS-90 Punch Card Configuration

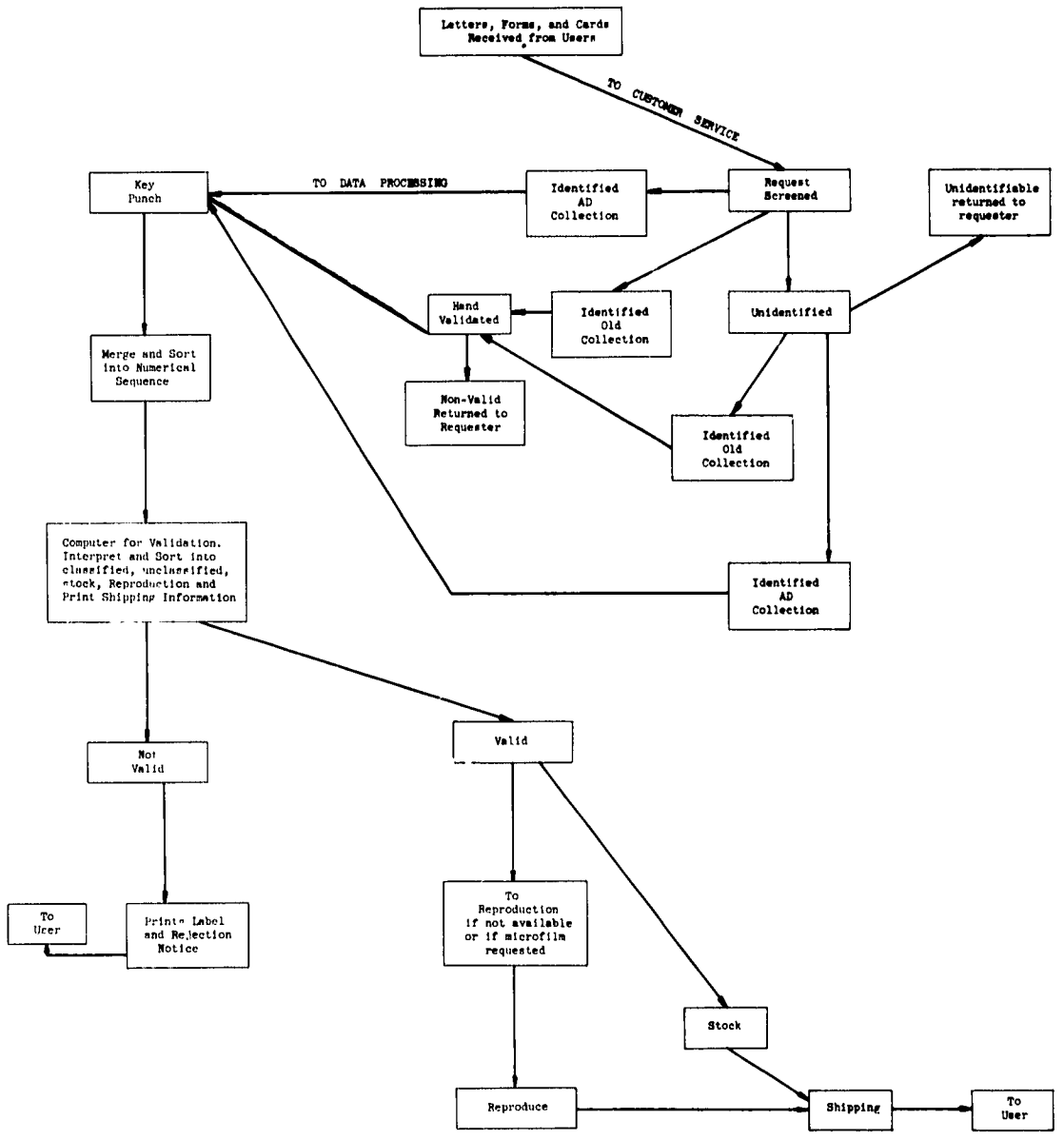


Figure 11

REQUEST PROCESSING
USS 90 Tape Configuration

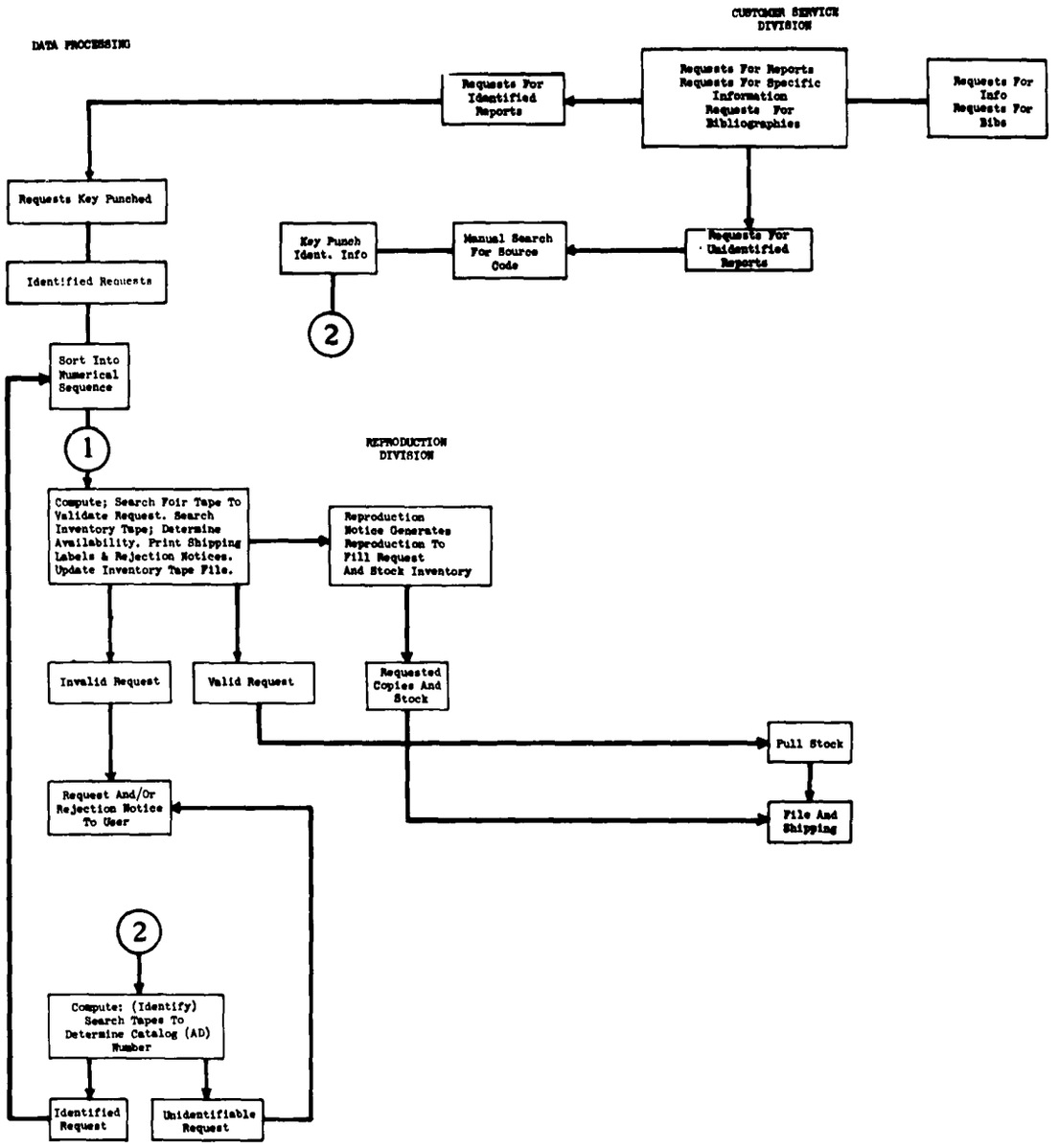
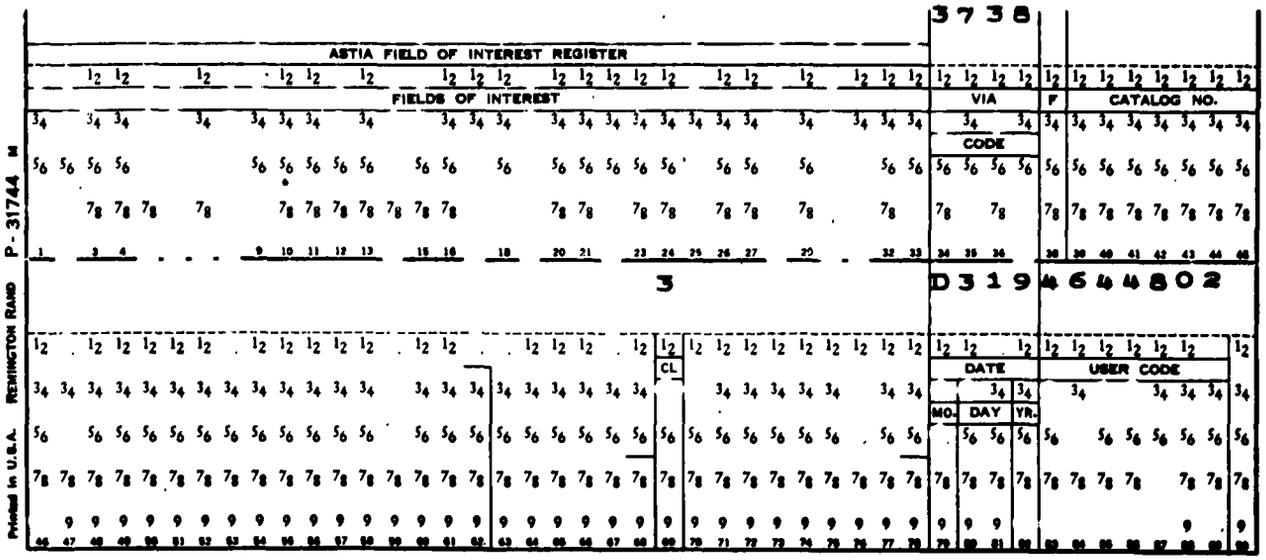


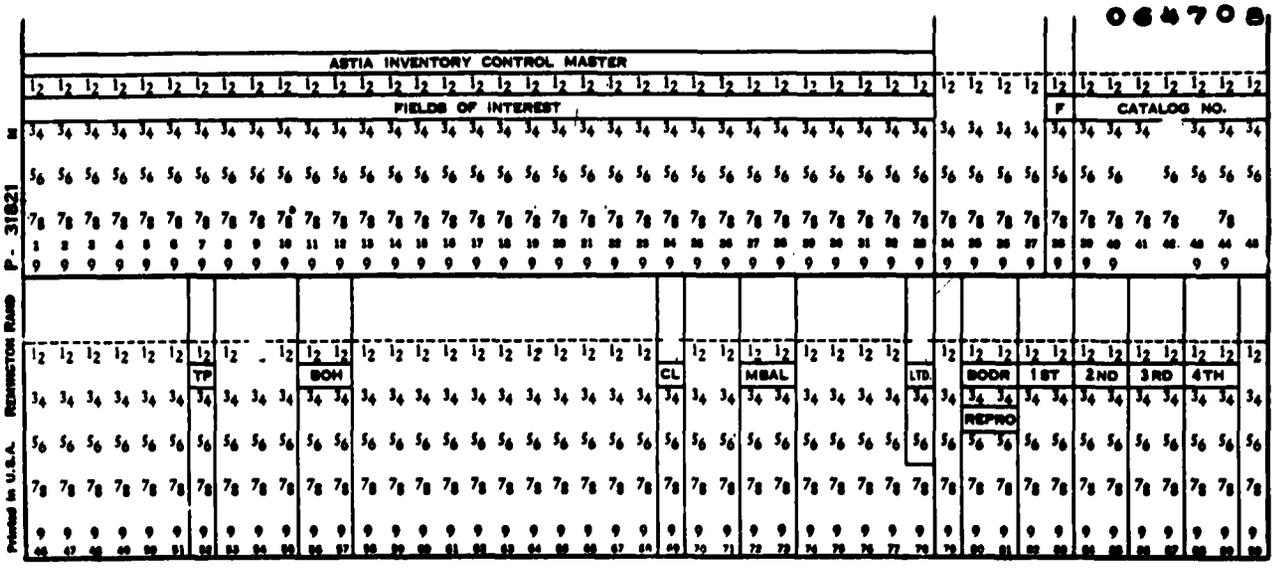
Figure 12



ASTIA Field of Interest Register

Information from ASTIA FOIR has been punched into card configurations shown above. Approximately 5,000 punched cards are required for this file. Approximately 3,000 cards applicable to contractors and 2,000 cards applicable to military users. This card format is also used as an internal processing request card. The card is retained as a record of the transaction for accumulation of statistical data.

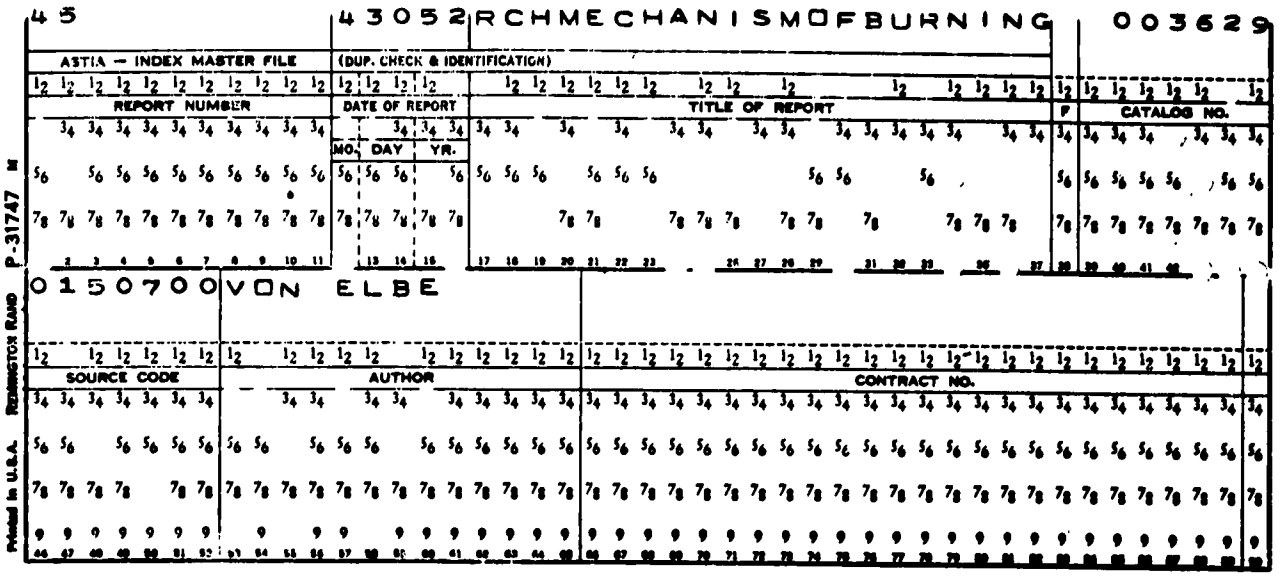
Figure 13



ASTIA Inventory Control Master File

Information for each document in the AD collection has been punched into card configuration shown above.

Figure 14

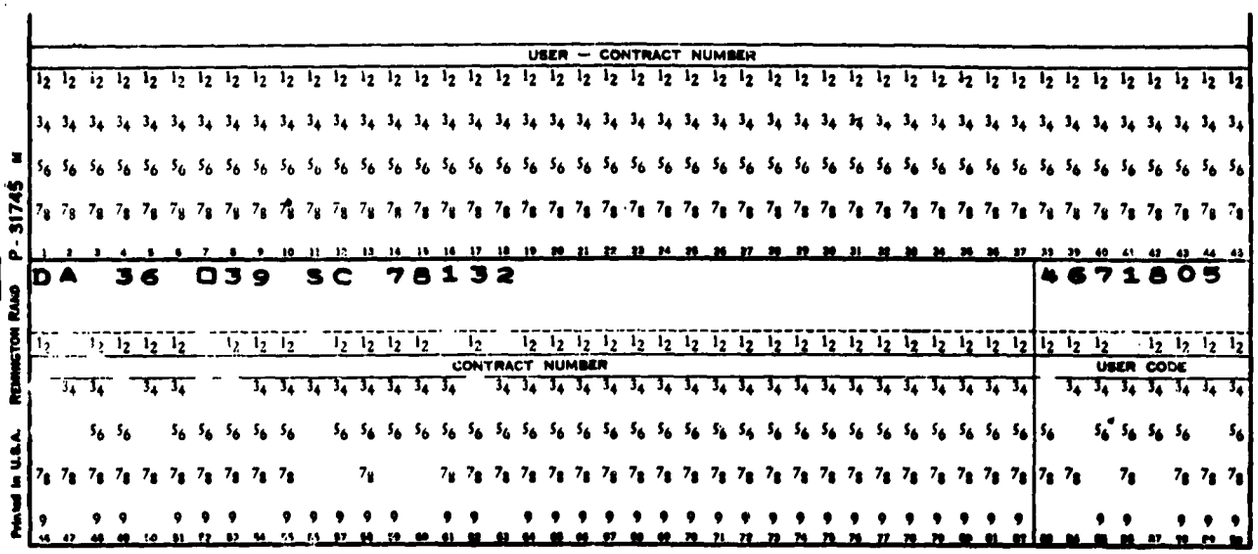


ASTIA Index Master File

This card file is being prepared for input to the magnetic tape system. Information from the ASTIA Catalog Cards is being key punched into the configuration shown above. Trailer cards are being prepared when more than one source is credited. A separate trailer card is also prepared when cross reference is desired to a military agency report number.

In the tape operation the data contained in the ASTIA Master Index File will be used to identify documents that are not requested by ASTIA catalog number and to check incoming documents to determine if they have previously been cataloged into the system.

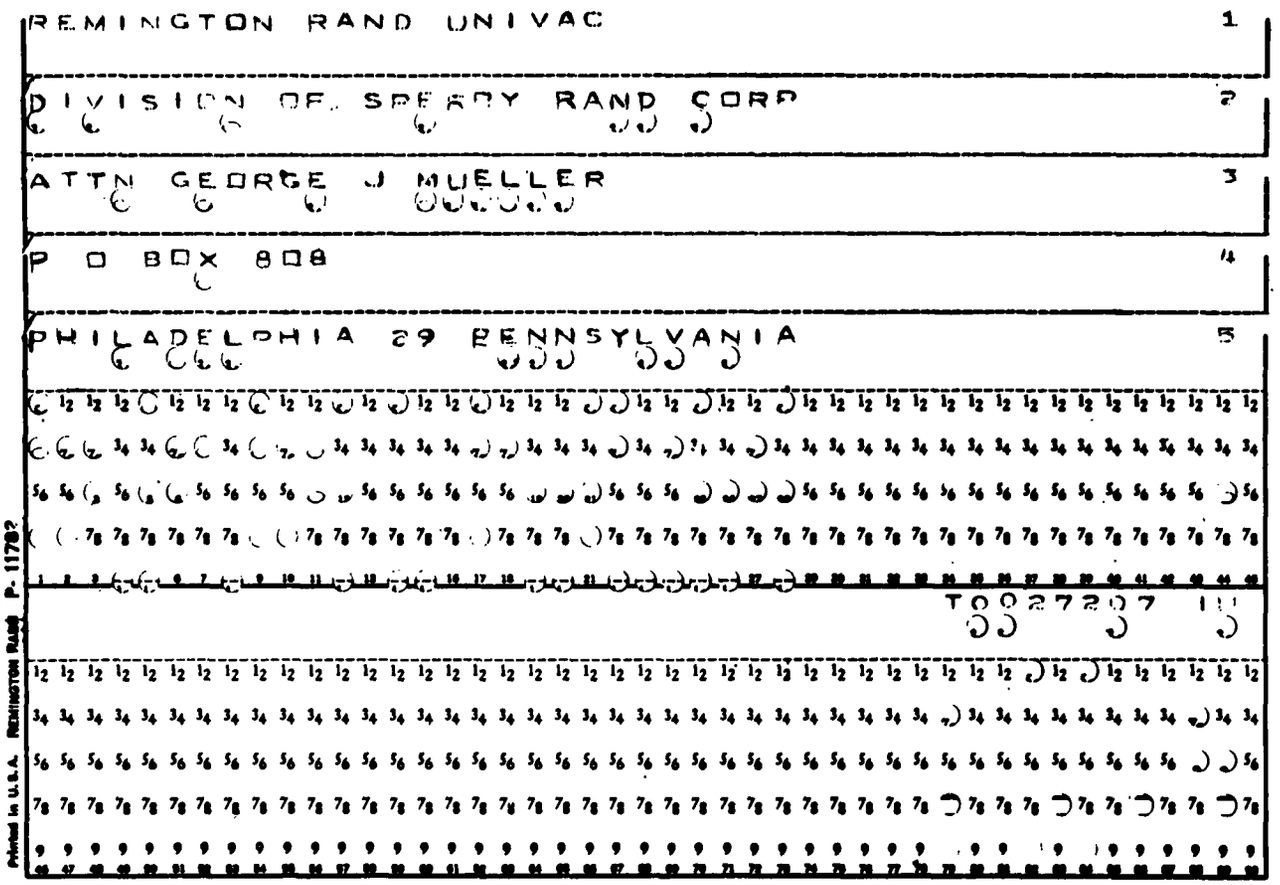
Figure 15



User Contract -- User Code File

This file is maintained both in user code sequence and in contract alpha-numeric sequence. It is not employed in the machine runs but is maintained for manual reference and record control. Copies will also be distributed to the appropriate contractor and via officer.

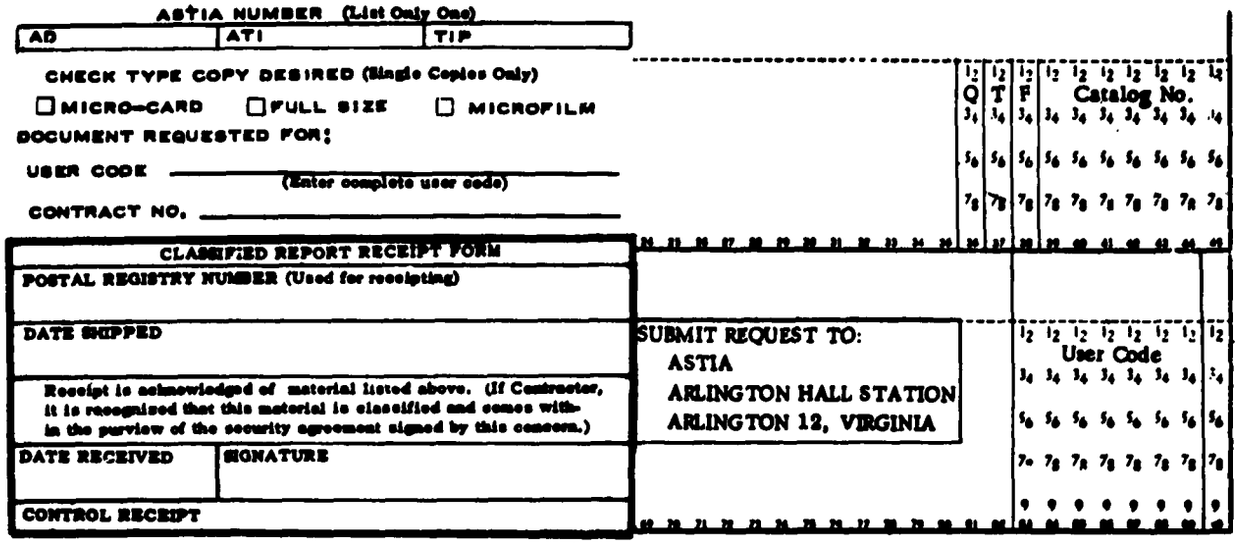
Figure 16



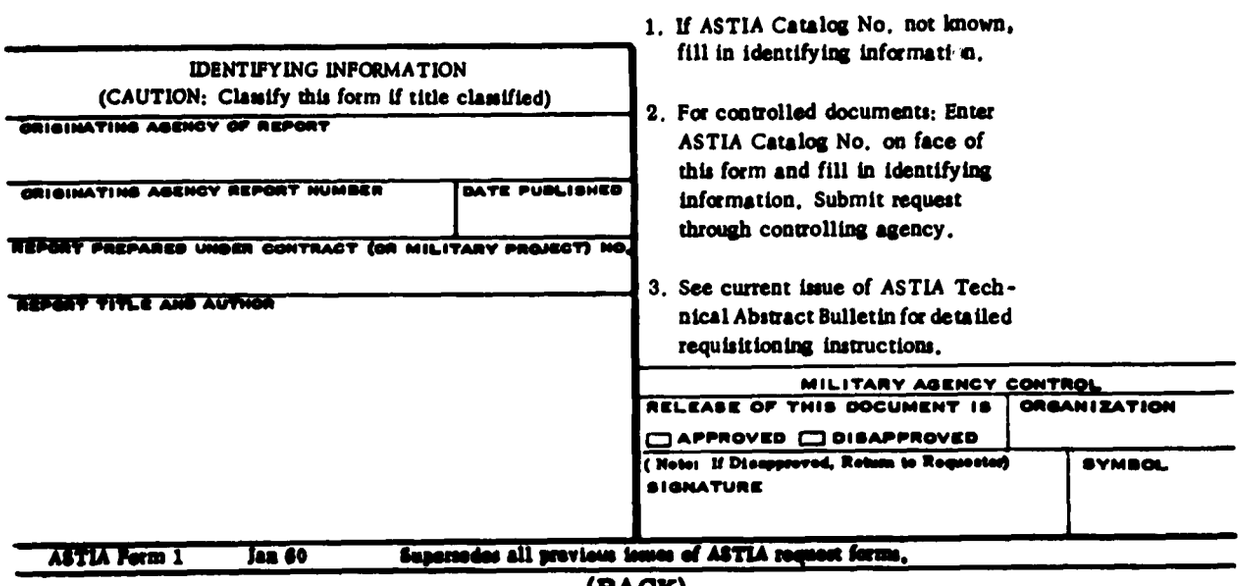
Address File

Five standard 90 column punched cards are used for each address. In the card configuration the address file will be maintained for record only. The data will be transferred to the tape configuration for print out of address on rejection notices and shipping information high-speed printer form letters.

Figure 17



(FRONT)



(BACK)

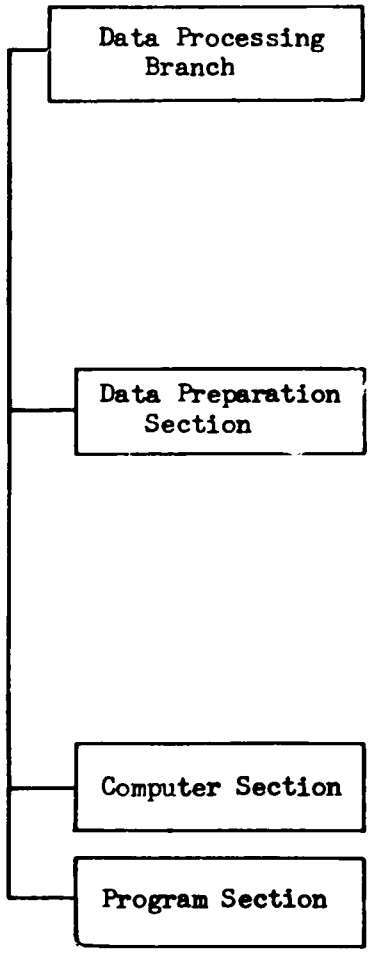
ASTIA Form 1

Punched-card request form, approximately one-half million requests will be submitted to ASTIA annually on this card. The card forms will be furnished each authorized ASTIA customer.

Figure 18

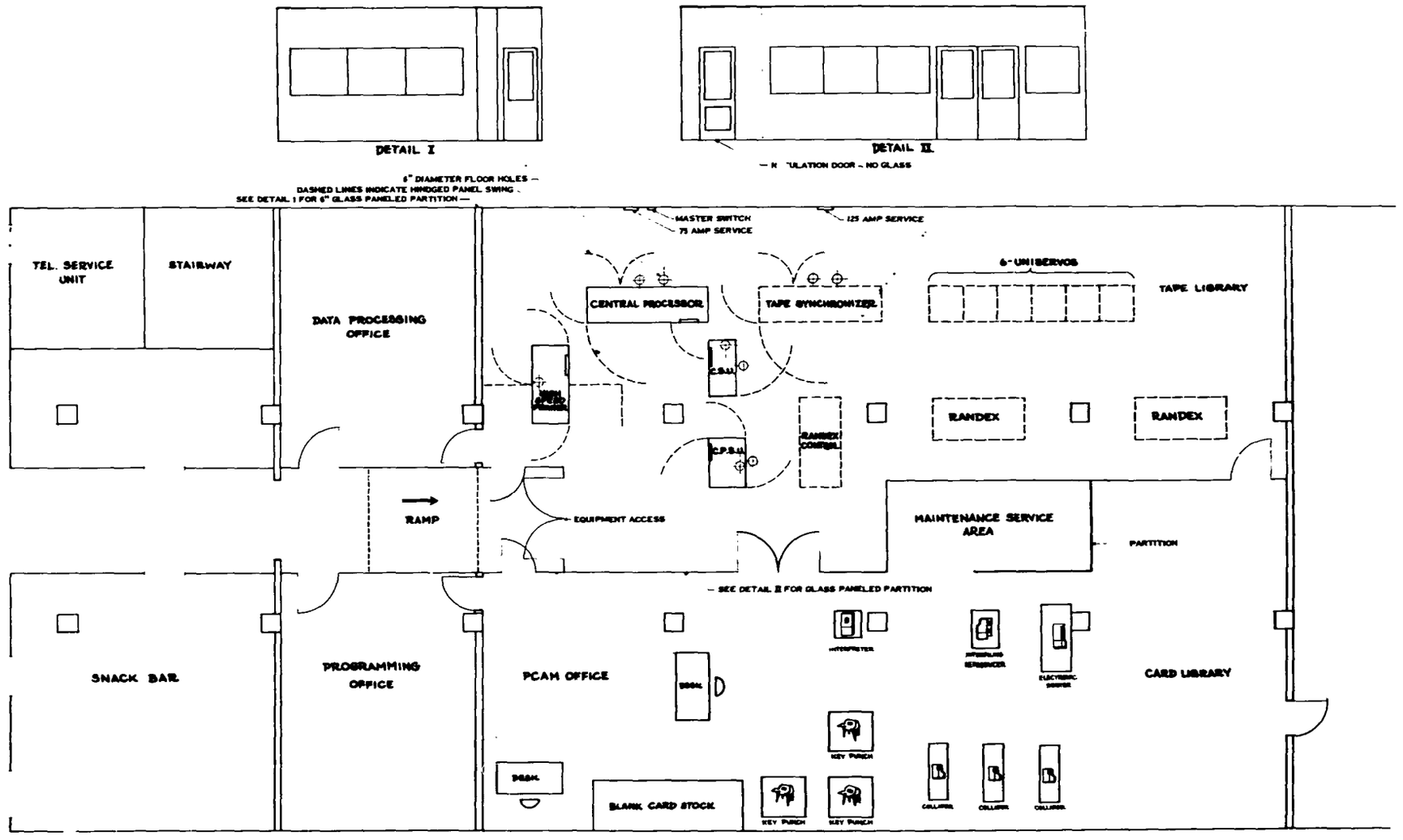
ORGANIZATION FOR DATA PROCESSING

Management Division



Organization	Function	Manning	Reporting Date
Data Processing Branch	1. Supervises Operation of ASTIA (1) Automatic Data Processing Facilities (3) 2. Directs Implementation of Automation of ASTIA Operations (2) in accordance with Approved Schedule 3. Administers and Schedules Operation 4. Maintains Liaison with Other Activities and Operations	1 - Chief, Data Processing Branch 1 - Deputy Chief and Procedure Supervisor 1 - Secretary 3 Total	Military on board 3 August 59 10 July 59
Data Preparation Section	1. Converts Data into Punched Card Form (5) 2. Controls Receipt and Release of Input/Output (6) 3. Operates Off-Line Punch Card Data Processing Machines, such as: Verifier, Collator, Tabulator, Sorter, and Reproducer (7) (8) 4. Maintains Active and Inactive Punch Card Files (21) (22) 5. Maintains Magnetic Tape Records	1 - Chief, Data Preparation Section (PCAM) 1 - Clerk-Typist/Key punch Operator 3 - Key Punch/Unityper Operators 2 - EAM Operators 1 - File Clerk (Card File) 1 - Library Assistant(Tape) 9 Total	20 July 59 27 July 59 1- 3 August 59 1- 30 November 59 1- 15 January 60 15 January 60 5 January 60 1 April 60
Computer Section	1. Operates Electronic Computer (11) and On-Line Peripheral Equipment (12)	1 - Chief, Computer Section 1 - Console Operator 2 Total	1 November 59 1 November 59
Program Section	1. Translates and Converts Operational Flow into Detailed Machine Instructions for the Computer (14) (15) (16) (18) 2. Develops Testing and Debugging Procedures for Programs 3. Establishes routines and Subroutines	1 - Chief Programmer 1 - Senior Programmer 1 - Junior Programmer 1 - Clerk Stenographer 4 Total	5 October 59 1 July 59 1 July 59 5 October 59
	TOTAL	18 (1 Military, 17 Civilians)	

Figure 19



WING 4 - BUILDING B

Figure 20.

40

NOTES:

1. ALL UNITS TO BE SUPPLIED WITH 240 V. 1ϕ - 3 WIRE
2. POWER REQUIREMENTS & HEAT DISSIPATION OF THE ABOVE UNITS IS AS FOLLOWS:
   - CENTRAL PROCESSOR - 5 KW. (7·1 KVA.) 284 BTU/MIN.
   - HIGH SPEED PRINTER - 3·5 KW. (4·8 KVA.) 198·5 BTU/MIN.
   - C.S.U. - 1 KW. (1·25 KVA) 94·6 BTU/MIN.
   - C.P.S.U. - 1 K.W. (1·25 KVA) 94·6 BTU/MIN.
   - TAPE SYNCHRONIZER - 6 KW. (7·5 KVA.) 370 BTU/MIN.
   - RANDEX CONTROL - 5·5 KW. (6·6 KVA) 313·5 BTU/MIN.
   - RANDEX - 2·2 KW (3·9 KVA) 155 BTU/MIN. PLUS ·5 KW. FROM CONTROL (INSTALLATION SHOWS 2-DRUMS)
   - UNISERVO - 2·2 KW (2·5 KVA) 124·3 BTU/MIN. (INSTALLATION SHOWS 6-UNISERVOS)

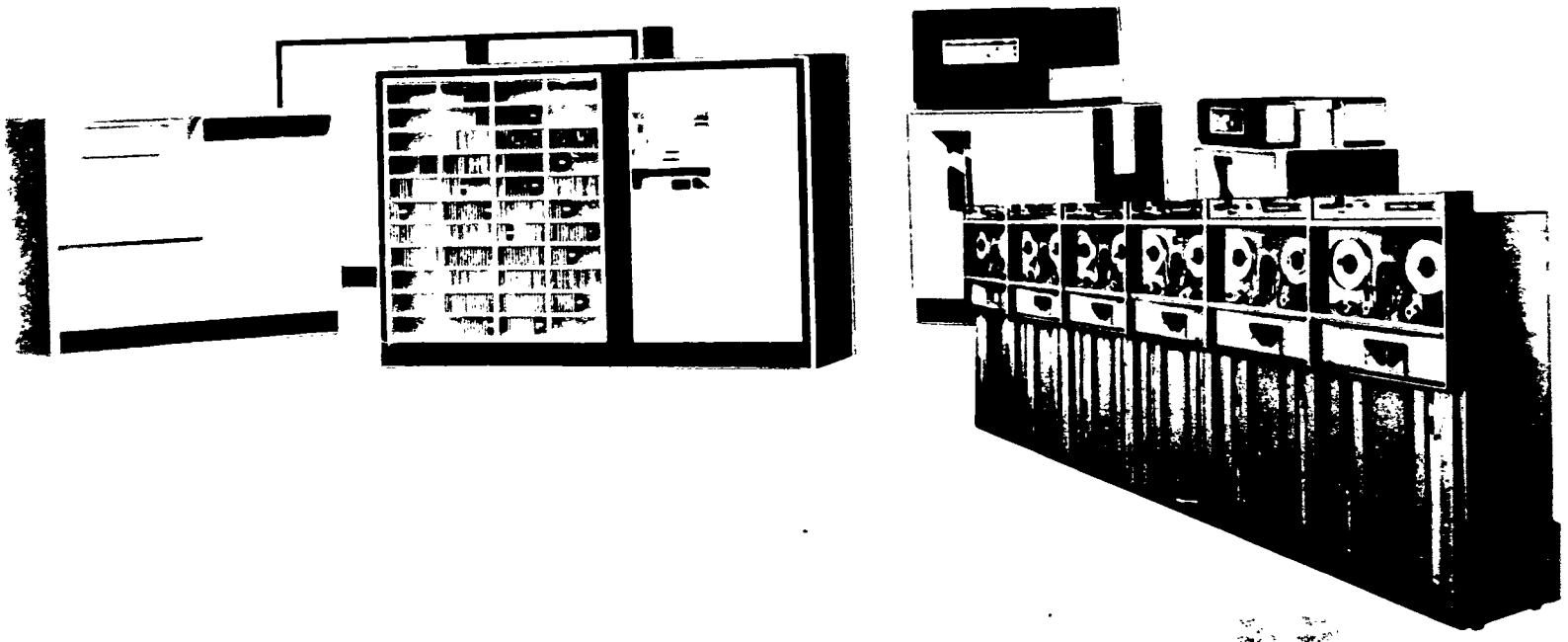


FIGURE 21

UNIVAC SOLID STATE 90

with 6 Magnetic Tape Uniservo II's

AUTOMATION OF ASTIA

SECTION III

CREATION OF A SYSTEM OF SCIENTIFIC MACHINE RETRIEVAL TERMS

(PROJECT MARS)

J. Heston Heald

Chief, Document Processing Division

As ASTIA began its plans for automation and began making decisions as to procedures and equipment, the most subtle of all problems was the development of a scheme for automatic subject control of its vast store of scientific information. How could a librarian, in subject-analyzing a scientific paper today, be sure he would assign the term which a scientist would use to call for the information, say, five or ten years hence? Of course, as long as the human element is involved at the input and the output ends, selection of the same term carries an element of chance. But perhaps this element could be reduced some degree at least. Certainly many keen minds had been devoted to this problem. Logicians, librarians, mathematicians, and the scientific community in general have been concerned. Numerous mechanical or even automatic schemes have been devised and a number of them actually put into operation, but none long enough to be considered safely proven. More important, none had ever been attempted in connection with a collection the size and scope held by ASTIA. In any case then, new ground would be broken. The stirring stimulus to move forward was that automation could never be reached and made workable without starting and giving it an honest try.

Most challenging was the fact that the manual system of subject-heading control which ASTIA had built over the years had a proven degree of success. The answer to this challenge must await many months of trial and experience and must be based not only on quality, thoroughness, and accuracy of retrieval, but also in terms of timeliness, cost, and personnel requirements. These factors will not be easy to weigh since the final yardstick will be in terms of value to scientific research and development.

For a number of years, ASTIA and its predecessor elements in Dayton and at the Library of Congress employed the previously mentioned "subject heading" scheme to provide subject matter access points to its report literature. These headings developed and the list grew as the literature entered the collection and was analyzed by the staff. It thus became an authoritative list; meaning that for each heading, information existed in the system.

By early 1959, the list of subject headings numbered over 70,000. These were published in the Fourth Edition, ASTIA Subject Headings. Also, in 1953, ASTIA began assigning Uniterms (single word headings) to each report and these along with the subject headings were printed on the catalog cards. Actually, it was with automation in mind that the Uniterm program was pursued. There was an apparent ease with which Uniterms could be manipulated into machine language and their retrieval accomplished by coordination or collation of related terms. Also, it was felt that Uniterms permitted much more liberal use than subject headings. The manual subject heading operation dictated a degree of conservative usage to keep it within both manageable and logical bounds.

As the realities of automation were eventually faced, it became obvious that neither the subject headings nor the Uniterms answered the requirements. Both had much to offer, but there was something yet to be desired.

It was decided that the "something" lacking was a Thesaurus-like arrangement of scientific and technical terms. This could provide a common guide to both the storage and the retrieval processes.

Actually, the list of subject headings was somewhat Thesaurus-like in structure. Because of this, and because they were authoritative, subject headings demanded consideration in the transition to any new system. But the rather complex arrangement with subdivisions and cross-references, strong for a manual operation, did not permit the full versatility of automatic data processing techniques. Also, the Subject Headings had a tendency to grow continually to accommodate the literature. There could be very little

leveling off. Obsolete or unneeded terms were never dropped. As for the Uniterms, they lacked definition; no over-all list had been maintained and very little consistency had been used in their assignment. Nevertheless, they were valuable in that they gave description and definition to the report to which they were assigned.

Based on these thoughts and the experience gained by actually using these schemes, the decision was made to establish an entirely new subject analysis concept for the ASTIA Machine Retrieval System. Hence, Project MARS was born.

Project MARS has two primary objectives:

1. To prepare a Thesaurus of scientific Descriptors, and

2. To assign these Descriptors to all AD-numbered reports in the ASTIA collection.

The Document Processing Division of ASTIA was assigned the responsibility for the Project and a group of leaders in the Division was named to work as a team. The group was composed of people with both scientific and library training and experience. Most important was the fact that each had been in the ASTIA program for a number of years and was quite familiar with the collections and the bibliographic problems. In addition to the writer, the members of the group include: Mr. Herbert Rehbock, Chief, Scientific Analysis Branch; Mr. Paul Klingbiel, Editor, Technical Abstract Bulletin (TAB) Mrs. Jane Virginia Philbrick, Assistant Editor, TAB; Mr. Martin Brooks, Chief, Engineering Section; Mr. Paul Klinefelter, Chief, Bio-Sciences Section; Mr. John Moats, Chief, Material Sciences Section; Mrs. Doretha Bebbs, Chief, Physical Sciences Section. This team represented a total of almost 100 years experience in the bibliographic organization of scientific and technical information.

Before a single move was made, the team spent several weeks reviewing essentially all known retrieval systems and exploring their possibilities with respect to the collections in ASTIA. American University arranged and gave a special course in this field. Several authorities were consulted, including Mrs. Lea Bohnert of RCA (leader of the course); Mr. Calvin Mooers, President, the Zator Company, who was employed at various intervals on a consultant basis; and Dr. John Mauchly, President, Mauchly Associates, who was also engaged to spend one three-day session with the group.

Decisions were reached through deliberations of the team and it met for short periods almost daily, often after the regular working day had ended. Each made contributions along the way and numerous schemes were proposed and tried before decisions were reached. Finally, it was decided to develop a Thesaurus of scientific terms based primarily on the ASTIA Subject Headings.

The first step, then, was an overhaul of the Subject Headings. The principal headings were divorced from their sub-divisions and, in one move, the list was reduced from 70,000 to about 8,300 main headings. The sub-divisions numbered another 850. Further steps in refinement eliminated about 150 of these sub-divisions as being no longer useful. The sub-division "application" is an example. This word had been tacked to main headings hundreds of times. The sub-divisions which were retained were either established as principal headings in their own right or included in other headings as synonymous or definitive terms. This first step, then, left us with a total of slightly more than 9,000 headings.

At this point, in order to differentiate from Subject Headings, the new headings were called "Descriptors." In many cases, they were not single words, although, combinations

of two and three (or more) words often remained. An example is "Radio frequency power." Hence, Descriptors became authentic "units of information" which could be used for analytical treatment of the documents. They were "authoritative" because they initially came from the literature in ASTIA, via the Subject Headings. The Subject Heading pattern of direct rather than inverse arrangement of terms was continued.

In order to define these Descriptors, scope notes were developed showing what the Descriptors included, what they were related to, or what they were limited to. This was done by listing synonyms, other Descriptors with which they were partially related, and areas which were or were not to be included in the meaning. Synonyms and the related or inclusive terms appear in their alphabetic positions referring the reader to the parent Descriptor. This treatment permitted further reduction in the number of Descriptors until the 9,000 became less than 7,000. A small section of Descriptors from the Thesaurus shows the order of arrangement:

ACARICIDES
(PEST CONTROL AND INHIBITING AGENTS)
INCL: MITICIDES
ALSO SEE: ANTIPEST IMPREGNANTS
PARATHION
PEST CONTROL

ACCELERATION
(MECHANICS)
ALSO SEE: DECELERATION

ACCELERATION INTEGRATORS; USE
ACCELEROMETERS

ACCELERATION TOLERANCE
(TOLERANCES)

ACCELERATORS
(PARTICLE ACCELERATORS)
ALSO SEE: BETATRONS
CYCLOTRONS
ELECTRON ACCELERATORS
ELECTROSTATIC ACCELERATORS
ION ACCELERATORS
LINEAR ACCELERATORS
PARTICLE ACCELERATORS
PROTON ACCELERATORS
SYNCHROTRONS

The terms on the far left margin are the main entries. The next indentation marks the beginning of scope notes for the entry. The parenthetic entry denotes the "schedule" of Descriptors to which this main entry belongs. "Incl." means <u>includes.</u> For example "Acaricides" <u>includes</u> "Miticides." "Acaricides" is a Descriptor, "Miticides" is not. However, "Miticides" will appear in the Thesaurus in its alphabetical order on the left margin, but there the words "use Acaricides" will appear. Note above that "Acceleration Integrators" refers the reader to <u>use</u> "Accelerometers."

The <u>see also's</u> are Descriptors in their own right and will appear in their respective alphabetical order throughout the list. <u>See also's</u> are closely related and help define the area of the principal heading.

It should be noted here that a feeling of gratification and accomplishment became evident among the staff as this effort progressed. Most of them had been architects of the Subject Headings and had grown to realize that renovation was long overdue. Such was almost impossible heretofore because the list was a key to a manually maintained card catalog which must be changed if the key was changed. It was no easy job and it was seldom attempted. Hence, the only way the list could move was to get larger. So, the opportunity to make a clean-up presented itself and the staff eagerly took advantage of it. They were encouraged with the thought that machine changes could take care of the future.

Although the building of the Thesaurus started with the Subject Headings, refinement and scope notes were developed by consulting all bibliographic information (including Uniterms) dealing with the reports in the AD collection and by the actual assignment of Descriptors to the respective documents. This building by practical application seemed essential. The staff was not concerned by changes that might be necessary on early-processed reports since all information would be programmed into the computer and corrections or updating could be a machine process. Hence, it was possible to develop the Thesaurus and do a certain amount of assignment of Descriptors at the same time; each process supporting the growth of the other.

We now come to the preparation of accompanying tools for use by both the subject analyst and the reference or retrieval operation. Although the Thesaurus is the principal guide, it was soon realized that certain condensed guidelines and checkpoints as companions were needed for quick, easy, and active consultation at both ends of the program of storage and retrieval. All this is aimed toward the goal of trying to bring together the person who put the information into the system and the person who goes to get it out. The closer they can be brought together in their thinking and their approach, the more accurate the retrieval system will become.

In this regard, it is necessary to say something about the ASTIA organizational arrangement, primarily on the input side. In the Scientific Analysis Branch, there are about 25 people with scientific backgrounds, and so distributed that they represent most of the scientific disciplines. The normal workload is thus divided among these people in subject analysis. For example, a mathematician gets the mathematics reports and the physicist gets the light, sound, heat reports, and so on. Because of this division of work, it was logical that the Descriptors in the Thesaurus be divided into a number of generic groups for display purposes. Hence, the full scope of the subject coverage is divided into about 290 major categories. Descriptors in the Thesaurus were then placed in the categories in which they were considered most pertinent. No Descriptor was repeated. Each was placed in a category by definition, thus distinctive display schedules were formed. Here is one of the display schedules. The schedule of TOLERANCES and the Descriptors that are to be used under that schedule.

(268) TOLERANCES

ACCELERATION TOLERANCE
ATROPINE TOLERANCE
GLUCOSE TOLERANCE
HEAT TOLERANCE
POISON TOLERANCE
POTASSIUM TOLERANCE
RADIATION TOLERANCE

We now refer back to the Thesaurus example of the Descriptor ACCELERATION TOLERANCES and the parenthetic word below it, TOLERANCES. TOLERANCES is the

group or the schedule into which it is associated. The number at the left, 268, is the number assigned to that particular schedule for location purposes as well as for easy reference. When a Descriptor is assigned, the schedule number to which it belongs will be shown if that Descriptor is to be used for cumulative index purposes. This depends upon the closeness of the Descriptor to the main subject of the report and is a matter of decision by the analyst.

This arrangement has a notable feature which is of considerable importance. It gives a key for cumulative indexes. In other words, accumulations can now be made by use of the display schedules under automation. None will be overweighted since cumulative indexes are to be prepared quarterly, each covering six issues of the Technical Abstract Bulletin. Both ASTIA and ASTIA users have had a definite requirement for cumulative indexes to TAB.

On the output side of the operation, reference personnel preparing bibliographies and answering specific reference inquiries should be guided in as nearly the same channels of reasoning as are the subject analysts on the input side, by use of the Thesaurus and these display schedules.

Although few in number, changes will be necessary in both Descriptors and schedules as time goes on and as information is programmed into the system. Under computer control, it is possible to readily up-date by either adding or deleting Descriptors or by combining descriptors with other schedules, or by even creating new display schedules as they become apparent. Versatility here is much greater than manual operations could possibly afford.

We now come to what are called "open-ended" terms. These terms either lack the analytical structure of Descriptors or have very little depth in terms of the amount of information to which they refer. They include project names, code names, code numbers, names or numbers of pieces of hardware or actual equipment designators and perhaps key or title-like words which at the time of assignment would not be considered of sufficient analytical nature to be treated as a Descriptor. These are called "open-ended terms" because they will not appear in the Thesaurus or in any of the display schedules. Although a list of these terms will be maintained in ASTIA, it is not planned at this time to publish them. Here again, by use of the computer, the density or relative importance of any of these terms can be determined at any time so that we may change them to Descriptors when justified. Such a term as "Satellite rendezvous" started out as only an open-ended term, pinpointing a relatively small area of information. However, as time passed, information in this area increased until the term became a Descriptor in its own right. Some examples of other types of open-ended terms are:

Project Michael

Saber jet (also F-86, F100, etc.)

Eureka 72

Special terroflex

These open-ended terms permit considerable freedom for establishing retrieval points where Thesaurus control is not considered essential. Another advantage is that it permits the use of classified terms, thus keeping the Thesaurus unclassified.

Counting both Descriptors and open-ended terms, it now appears that there will be an

average assignment of about eight subject access points to each report in the AD numbered collection of ASTIA. For convenience, Descriptors and open-ended terms together are referred to as "retrieval terms."

Still another tool to help keep both analyst and searcher on the same track is the check list mentioned earlier.

As a document is processed the analyst checks over the list to be sure he has covered the pertinent elements with retrieval term assignment. Has he covered the subject with one or perhaps more Descriptors? Is it possible to assign a term or terms to show the trend of the research? Can terms be assigned which reflect something about the results? Was equipment used that should be picked up, perhaps as an open-ended term? Etc., etc. Again, this list may be used by reference people as an additional guide. An analysis of a reference question in terms of these elements will surely assist in selecting the pertinent Descriptors for retrieval purposes. The actual check list used follows:

CHECK LIST FOR ASSIGNMENT OF RETRIEVAL TERMS

1. Subject. (What was studied, investigated, tested, compiled, researched?)

2. How Was the Subject Treated? (analysis, tests, design, production, computations, theory, specifications, operations, processes)

3. What Are the Physical Factors? (mechanical properties, physical properties, chemical effects, biological factors)

4. What Equipment or Method Was Used to Support the Research or Investigation? (spectrum analyzers, oscilloscopes, Charpy v-notch test equipment)

5. Where or Under What Environment Was the Research Accomplished? (upper atmosphere, arctic, sub-surface, location if geographic or foreign)

6. Additional Qualifying Information (Open-ended Terms). (project names, military symbols, trade names, Mark/Mod numbers, AN/numbers, etc.)

To exemplify the use of this system under automation, a sample question is: "What information does ASTIA have concerning the accuracy of inertial systems for use in Air-to-surface bombing by guided missiles?"

The next step would be to consult the Thesaurus for pertinent terms. The following Descriptors listed appear pertinent:

1. Guided missiles

2. Guidance

3. Air-to-surface

4. Inertial guidance

5. Bombing systems

Obviously, if the term "Guided missiles" alone is used, the retrieval would include much unwanted information. In fact, there are about 13,000 reports in the AD numbered collection of ASTIA dealing with "Guided missiles." Correlation with the other

Descriptors is then necessary and can be done by computer. Adding the Descriptor "Guidance" reduces the field to some 3,000 reports. Each additional Descriptor added continues to narrow the output until finally redundant information is essentially eliminated and, by coordinating all of the above Descriptors, a total of 12 reports, all related to the question above is found. An open-ended term related to this subject could be the name or designator of the system. If the inquirer knows the name or designator of this particular system, he can then retrieve quite specifically the information he wants. Of course, if he doesn't know the name, information concerning it will be retrieved along with all similar systems under the first procedure.

In order to avoid use of terms which might not have been used by the analyst on the input operation, the use of the Thesaurus, the displays, and the check list are important.

In summary, this has only in a brief way described the development of ASTIA's transition from a manual system using Subject Headings, with its sub-divisions and cross-references, into a subject arrangement of the collections tailored for computer programming and retrieval. The Uniterms that were assigned in the past have been used primarily as aids in defining the Descriptors although in some cases they will actually become Descriptors themselves. The Descriptors are arranged into two arrays:

(1) An alphabetical listing in the Thesaurus where each has its scope notes or definitions as appropriate and the major display field into which it has been categorized.

(2) Into the display groups.

As this is written, the program is well under way. The Thesaurus is not yet ready for publication nor may it be expected until sometime after the first of the year. It is planned to be operational under this retrieval system by 30 June 1960.

A concentrated and time-consuming effort is being applied to the fundamental problem of adapting ASTIA's available scientific terminology to the specific needs of machine retrieval. Important phases of this reappraisal include: the elimination of redundancy in meaning of the terminology; minimization of foreseeable false correlations through combination and permutation of retrieval terms; proper breakdown of retrieval term groups to assure adequate depth of retrieval; and simplification and coordination of terms in the framework of the major scientific disciplines.

Our primary objectives for this information retrieval phase of our program are to:

(1) Reduce the time required to prepare bibliographies and answer reference inquiries.

(2) Improve the quality of ASTIA's information services.

(3) Automatically prepare periodic cumulative indexes to TAB, as well as cumulations in any given chronological period and subject area that may be desired.

(4) Eliminate the need for the time-consuming and costly tasks of catalog maintenance. In ASTIA this amounts to almost $100,000 annually in personal services alone.

(5) Provide desk-copy reference tools, in terms of TAB, the cumulative indexes, and the Thesaurus, to the librarians, the scientists, the engineers, and the project officers in the research and development program; such tools to eventually become tantamount to card catalogs as an approach to the ASTIA collections.

AD-227 000 Div. 32. 30

Armed Services Technical Information Agency, Arlington, Va.
AUTOMATION OF ASTIA, by W. A. Barden, William Hammond, and J. Heston Heald. Preliminary rept. 1 Dec 59, 50p. incl. illus. Unclassified report

Early considerations in automation: The history of ASTIA's experience in planning and implementing the automation of its functions is presented. Different ideas were examined and discarded in a search for a more efficient method of indexing and retrieving information. In 1953 a preliminary study based on a systems concept embracing all the functions and services of the Agency was conducted, but a full scale study was not possible until 1958. When the final selection of the Remington Rand USS-90 (Univac Solid State Computer) was made, the ASTIA staff devised methods for making optimum use of the equipment in both the business-type and information retrieval functions. Automation program: The pre-automation and automated processing (over)

UNCLASSIFIED

Barden, W. A.
Hammond, William
Heald, J. Heston

RETRIEVAL TERMS

Data processing systems*
Programming*
Punched card methods*
Scientific reports*
Documentation*
Digital computers*
Military research*
Univac
Randex
Project MARS

Armed Services Technical Information Agency
UNCLASSIFIED

AD-227 000 Div. 32, 30

Armed Services Technical Information Agency, Arlington, Va.
AUTOMATION OF ASTIA, by W. A. Barden, William Hammond, and J. Heston Heald. Preliminary rept. 1 Dec 59, 50p. incl. illus. Unclassified report

Early considerations in automation: The history of ASTIA's experience in planning and implementing the automation of its functions is presented. Different ideas were examined and discarded in a search for a more efficient method of indexing and retrieving information. In 1953 a preliminary study based on a systems concept embracing all the functions and services of the Agency was conducted, but a full scale study was not possible until 1958. When the final selection of the Remington Rand USS-90 (Univac Solid State Computer) was made, the ASTIA staff devised methods for making optimum use of the equipment in both the business-type and information retrieval functions. Automation program: The pre-automation and automated processing (over)

UNCLASSIFIED

Barden, W. A.
Hammond, William
Heald, J. Heston

RETRIEVAL TERMS

Data processing systems*
Programming*
Punched card methods*
Scientific reports*
Documentation*
Digital computers*
Military research*
Univac
Randex
Project MARS

Armed Services Technical Information Agency
UNCLASSIFIED

AD-227 000 Div. 32. 30

Armed Services Technical Information Agency, Arlington, Va.
AUTOMATION OF ASTIA, by W. A. Barden, William Hammond, and J. Heston Heald. Preliminary rept. 1 Dec 59, 50p. incl. illus. Unclassified report

Early considerations in automation: The history of ASTIA's experience in planning and implementing the automation of its functions is presented. Different ideas were examined and discarded in a search for a more efficient method of indexing and retrieving information. In 1953 a preliminary study based on a systems concept embracing all the functions and services of the Agency was conducted, but a full scale study was not possible until 1958. When the final selection of the Remington Rand USS-90 (Univac Solid State Computer) was made, the ASTIA staff devised methods for making optimum use of the equipment in both the business-type and information retrieval functions. Automation program: The pre-automation and automated processing (over)

UNCLASSIFIED

Barden, W. A.
Hammond, William
Heald, J. Heston

RETRIEVAL TERMS

Data processing systems*
Programming*
Punched card methods*
Scientific reports*
Documentation*
Digital computers*
Military research*
Univac
Randex
Project MARS

Armed Services Technical Information Agency
UNCLASSIFIED

AD-227 000 Div. 32. 30

Armed Services Technical Information Agency, Arlington, Va.
AUTOMATION OF ASTIA, by W. A. Barden, William Hammond, and J. Heston Heald. Preliminary rept. 1 Dec 59, 50p. incl. illus. Unclassified report

Early considerations in automation: The history of ASTIA's experience in planning and implementing the automation of its functions is presented. Different ideas were examined and discarded in a search for a more efficient method of indexing and retrieving information. In 1953 a preliminary study based on a systems concept embracing all the functions and services of the Agency was conducted, but a full scale study was not possible until 1958. When the final selection of the Remington Rand USS-90 (Univac Solid State Computer) was made, the ASTIA staff devised methods for making optimum use of the equipment in both the business-type and information retrieval functions. Automation program: The pre-automation and automated processing (over)

UNCLASSIFIED

Barden, W. A.
Hammond, William
Heald, J. Heston

RETRIEVAL TERMS

Data processing systems*
Programming*
Punched card methods*
Scientific reports*
Documentation*
Digital computers*
Military research*
Univac
Randex
Project MARS

Armed Services Technical Information Agency
UNCLASSIFIED

AD-227 000

UNCLASSIFIED

of reports through ASTIA and validation of requests of military contractors is described. The three stages by which the automatic data processing system will be put into operation are examined, and the process of compiling mechanized cumulative indexes to the Technical Abstract Bulletin is presented. Creation of a Thesaurus of Scientific Descriptors: The main objectives of Project MARS (Machine Retrieval System) are: (1) to prepare a Thesaurus of descriptors; and (2) to assign these descriptors to all AD numbered reports in the ASTIA collection. The ASTIA subject headings and subdivisions were overhauled and the list reduced from 70,000 to about 9,000 headings, now termed descriptors. The scope of subject coverage was divided into about 290 generic categories called display schedules. Procedures were established for the assignment of retrieval terms, both standard descriptors from the Thesaurus and "open-ended terms" which will not appear in the Thesaurus but will provide additional retrieval access points in the form of project names, equipment nomenclature, trade names, etc.

UNCLASSIFIED

AD-227 000

UNCLASSIFIED

of reports through ASTIA and validation of requests of military contractors is described. The three stages by which the automatic data processing system will be put into operation are examined, and the process of compiling mechanized cumulative indexes to the Technical Abstract Bulletin is presented. Creation of a Thesaurus of Scientific Descriptors: The main objectives of Project MARS (Machine Retrieval System) are: (1) to prepare a Thesaurus of descriptors; and (2) to assign these descriptors to all AD numbered reports in the ASTIA collection. The ASTIA subject headings and subdivisions were overhauled and the list reduced from 70,000 to about 9,000 headings, now termed descriptors. The scope of subject coverage was divided into about 290 generic categories called display schedules. Procedures were established for the assignment of retrieval terms, both standard descriptors from the Thesaurus and "open-ended terms" which will not appear in the Thesaurus but will provide additional retrieval access points in the form of project names, equipment nomenclature, trade names, etc.

UNCLASSIFIED

AD-227 000

UNCLASSIFIED

of reports through ASTIA and validation of requests of military contractors is described. The three stages by which the automatic data processing system will be put into operation are examined, and the process of compiling mechanized cumulative indexes to the Technical Abstract Bulletin is presented. Creation of a Thesaurus of Scientific Descriptors: The main objectives of Project MARS (Machine Retrieval System) are: (1) to prepare a Thesaurus of descriptors; and (2) to assign these descriptors to all AD numbered reports in the ASTIA collection. The ASTIA subject headings and subdivisions were overhauled and the list reduced from 70,000 to about 9,000 headings, now termed descriptors. The scope of subject coverage was divided into about 290 generic categories called display schedules. Procedures were established for the assignment of retrieval terms, both standard descriptors from the Thesaurus and "open-ended terms" which will not appear in the Thesaurus but will provide additional retrieval access points in the form of project names, equipment nomenclature, trade names, etc.

UNCLASSIFIED

AD-227 000

UNCLASSIFIED

of reports through ASTIA and validation of requests of military contractors is described. The three stages by which the automatic data processing system will be put into operation are examined, and the process of compiling mechanized cumulative indexes to the Technical Abstract Bulletin is presented. Creation of a Thesaurus of Scientific Descriptors: The main objectives of Project MARS (Machine Retrieval System) are: (1) to prepare a Thesaurus of descriptors; and (2) to assign these descriptors to all AD numbered reports in the ASTIA collection. The ASTIA subject headings and subdivisions were overhauled and the list reduced from 70,000 to about 9,000 headings, now termed descriptors. The scope of subject coverage was divided into about 290 generic categories called display schedules. Procedures were established for the assignment of retrieval terms, both standard descriptors from the Thesaurus and "open-ended terms" which will not appear in the Thesaurus but will provide additional retrieval access points in the form of project names, equipment nomenclature, trade names, etc.

UNCLASSIFIED